

<https://helda.helsinki.fi>

Aspects of capacity of entanglement

de Boer, Jan

2019-03-27

de Boer , J , Järvelä , J & Keski-Vakkuri , E 2019 , ' Aspects of capacity of entanglement ' ,
Physical Review D , vol. 99 , no. 6 , 066012 . <https://doi.org/10.1103/PhysRevD.99.066012>

<http://hdl.handle.net/10138/301219>

<https://doi.org/10.1103/PhysRevD.99.066012>

cc_by

publishedVersion

Downloaded from Helda, University of Helsinki institutional repository.

This is an electronic reprint of the original article.

This reprint may differ from the original in pagination and typographic detail.

Please cite the original version.

Aspects of capacity of entanglement

Jan de Boer,^{1,*} Jarkko Järvelä,^{2,3,†} and Esko Keski-Vakkuri^{2,3,‡}

¹*Institute for Theoretical Physics, University of Amsterdam, Science Park 904,
1098 XH Amsterdam, Netherlands*

²*Department of Physics, University of Helsinki, P.O. Box 64, FIN-00014, Finland*

³*Helsinki Institute of Physics, University of Helsinki, P.O. Box 64, FIN-00014, Finland*



(Received 29 November 2018; published 27 March 2019)

Many quantum information theoretic quantities are similar to and/or inspired by thermodynamic quantities, with entanglement entropy being a well-known example. In this paper, we study a less well-known example, capacity of entanglement, which is the quantum information theoretic counterpart of heat capacity. It can be defined as the second cumulant of the entanglement spectrum and can be loosely thought of as the variance in the entanglement entropy. We review the definition of capacity of entanglement and its relation to various other quantities such as fidelity susceptibility and Fisher information. We then calculate the capacity of entanglement for various quantum systems, conformal and nonconformal quantum field theories in various dimensions, and examine their holographic gravity duals. Resembling the relation between response coefficients and order parameter fluctuations in Landau-Ginzburg theories, the capacity of entanglement in field theory is related to integrated gravity fluctuations in the bulk. We address the question of measurability, in the context of proposals to measure entanglement and Rényi entropies by relating them to $U(1)$ charges fluctuating in and out of a subregion, for systems equivalent to non-interacting fermions. From our analysis, we find universal features in conformal field theories, in particular the area dependence of the capacity of entanglement appears to track that of the entanglement entropy. This relation is seen to be modified under perturbations from conformal invariance. In quenched $1+1$ dimensional CFTs, we compute the rate of growth of the capacity of entanglement. The result may be used to refine the interpretation of entanglement spreading being carried by ballistic propagation of entangled quasiparticle pairs created at the quench.

DOI: [10.1103/PhysRevD.99.066012](https://doi.org/10.1103/PhysRevD.99.066012)

I. INTRODUCTION

Entanglement entropy has been proven quite useful as a diagnostic of topological properties of ground states of quantum many-body systems, for a recent review see e.g., [1]. At the same time, starting with [2] it has also been playing an instrumental role in understanding the connection between the geometry of space-time and strongly coupled field theories in the AdS/CFT correspondence, a recent extensive review is [3].

In this paper, we study another quantity associated to a reduced density matrix, namely the capacity of entanglement. It is defined in the same way as one defines heat capacity for thermal systems, which was the original

motivation to introduce this quantity in [4] (see also [5]). Explicitly, if λ_i are the eigenvalues of a reduced density matrix, entanglement entropy is defined as $S_{EE} = -\sum \lambda_i \log \lambda_i$, and capacity of entanglement can be written as $C_E = \sum \lambda_i \log^2 \lambda_i - S_{EE}^2$. The latter can also be thought of as the variance of the distribution of $-\log \lambda_i$ with probability λ_i , and it is clear that it contains information about the width of the eigenvalue distribution of the reduced density matrix.

Our main motivation was to understand the connection between capacity of entanglement and quantum fluctuations, and to e.g., see whether capacity of entanglement sheds any interesting light on the accuracy of the semiclassical approximation in gravity. This connection turns out to be the following. As is well known, the Ryu-Takanagi formula [2] relates the entanglement entropy in a quantum field theory to the volume of a minimal surface (“RT surface”) in a dual anti-de Sitter spacetime. If we include quantum gravity fluctuations of the RT surface, and integrate over them, the integral does indeed compute the capacity of entanglement in the quantum field theory [6]. Thus, capacity of entanglement associates a “width” to

*J.deBoer@uva.nl

†jarkko.jarvela@helsinki.fi

‡esko.keski-vakkuri@helsinki.fi

Published by the American Physical Society under the terms of the [Creative Commons Attribution 4.0 International](https://creativecommons.org/licenses/by/4.0/) license. Further distribution of this work must maintain attribution to the author(s) and the published article's title, journal citation, and DOI. Funded by SCOAP³.

the entanglement entropy: a measure of quantum gravitational effects. This relationship is a leading contribution to a more complete sum of gravitational fluctuations. In [7], the RT surface was replaced by a domain wall with tension (a “cosmic brane”), including its complete gravitational backreaction. The volume of the brane in the backreacted geometry was shown to be equal to a variant of the Rényi entropies in the quantum field theory, coined “modular entropy” in [8]. This result was used in [6], independently the modular entropy was introduced and used to define the capacity of entanglement in the earlier work [4,5].

Entanglement entropy and capacity of entanglement are two features of the full entanglement spectrum [9] (the set of eigenvalues of the reduced density matrix). There has been growing interest to study the entanglement spectrum as a diagnostic of the phase structure of various systems. For our purposes, of special relevance are the studies of the entanglement spectrum for $1+1$ dimensional gapless [10–12] and gapped [13] systems. Originally it was also hoped that there would be a precise relation between the entanglement spectrum and the energy eigenvalue spectrum, but there are caveats [14]. Connections between the capacity of entanglement and the entanglement spectrum have been explored in [15].

Complete knowledge of all the Rényi entropies is in principle sufficient to recover the complete set of eigenvalues of the reduced density matrix, the entanglement spectrum distribution, by a suitable integral transform. In this way, [10] derived the spectrum for $1+1$ conformal field theories in the ground state, with a finite length ℓ interval as the subsystem, and found that the spectrum is (under some assumptions) characterized by a universal function, depending only on two parameters: the central charge c of the theory and the largest eigenvalue λ_{\max} . Universality was also found in gapped systems [13], while the assumptions of [10] were further studied in [16] and the spectrum was interpreted as a reparametrization of the Cardy formula counting energy levels in a CFT.

While we will make a few remarks about the full entanglement spectrum, we will mostly restrict to the two lowest moments or cumulants of the spectrum, which are precisely the entanglement entropy and capacity of entanglement. In terms of the modular Hamiltonian $K_A = -\log \rho_A$ they are the expectation value and variance of K_A respectively. As stated above, for systems of with a gravity dual, higher cumulants capture the whole series of the quantum gravitational fluctuations about the RT surface, giving the complete entanglement spectrum or the Rényi/modular entropies.

Besides looking at gravitational fluctuations, we will in this work study the capacity of entanglement from many different points of view. We develop a more comprehensive list of its properties and relations to other concepts of quantum information, not just in the context of holographic gauge/gravity duality. We study a range of quantum systems

from simple qubit systems and random pure states to spin chains, conformal and nonconformal quantum field theories, with and without a gravity dual, and also nonequilibrium quenched systems. We are especially interested in identifying universal features. For example, we find evidence for an “area law” for the capacity of entanglement in conformal theories. Curiously, in four-dimensional conformal field theories, in a fairly natural regularization scheme, the ratio between the coefficients in front of the area term in entanglement entropy and capacity of entanglement turns out to be precisely a/c , the ratio of the a and c anomaly coefficients, for spherical entangling surfaces. This area law suggests that most of the quantum fluctuations of the RT surface are located near the boundary of AdS and this therefore does not seem to shed much light on the size of local bulk quantum fluctuations.

Another important observation is that systems where all entanglement is carried by EPR pairs (pairs of qubits that contribute $\log 2$ to the entanglement entropy) have zero capacity of entanglement. Therefore, whenever we find that entanglement entropy and capacity of entanglement are approximately equal to each other, EPR pairs are not a very good approximation of the quantum state. As we will show, randomly entangled pairs of qubits give a much better description, and this can be used to e.g., sharpen the picture of ballistic propagation of entanglement carried by quasi-particles created at a quench.

The outline of this paper is as follows. Section II begins with some basic definitions and relations. We begin with definitions of entropy, capacity of entanglement, and modular entropy based on quantum information theory. We discuss the moments and cumulants of the entanglement spectrum/modular Hamiltonian, and show how to obtain them from the Rényi entropies. We then move to analogues of thermodynamic definitions of entropy and heat capacity applied to quantum entanglement, and show agreement with the previous quantum information theoretic definitions. We move to review the definitions of fidelity susceptibility and quantum Fisher information, and show how they are related to the capacity of entanglement (related discussion is also in [17,18]). Finally, we map capacity of entanglement to the heat capacity of a thermal CFT on a hyperbolic sphere, recovering the thermodynamic equality of heat capacity and variance of entropy.

Section III discusses some properties of the capacity of entanglement. We first derive an upper bound for it in quantum systems with a finite dimensional Hilbert space. We also recast this result as an estimate of the variance in classical information theory. Addressing the measurability of the capacity of entanglement, we follow the proposal to relate bipartite entanglement for noninteracting fermions with fluctuations of conserved $U(1)$ charges [19–23]. We show how the moments and cumulants of entanglement spectrum and $U(1)$ charges are related. We point out that an equality between the entanglement entropy and capacity of

entanglement arises naturally in this context in a large N limit, at the leading level. We speculate how such a limit might arise in theories with a holographic bulk gravity dual, and consider an example of an eigenvalue distribution giving rise to the equality $C_E = S_{EE}$. We then interpret the capacity of entanglement holographically. We give a shorter derivation of the bulk integral formula of [6], and discuss its interpretation.

In Sec. IV, we study how the capacity of entanglement depends on the initial quantum state of the complete system. In simple n -qubit examples, the entanglement entropy and the capacity of entanglement are in general not proportional. For subsystems of random pure states, the approximate equality of the two can arise as ensemble averages, and we compute the precise ensemble averages for several choices of the dimension of the parent Hilbert space where the pure states are defined and choices of dimension of the subsystem.

In Sec. V, we study the capacity of entanglement in field theories. First we calculate it in $1+1$ CFTs at equilibrium and in local and global quenches. In all cases, the capacity of entanglement is equal to the entanglement entropy. For the nonequilibrium cases this implies that the rate of growth of the capacity of entanglement follows that of S_{EE} . We comment on the interpretation of ballistic spreading of entanglement. We then extend the computation to higher dimensions, following an alternative method based on stress-tensor correlators [24,25]. After that, we study universality. We discuss some aspects of the entanglement spectrum and its (non)universality. Then we consider the possible dependence on regularization schemes. In generic quantum field theories and CFTs, the ratio C_E/S_{EE} depends on the regularization scheme, but for CFTs with a holographic gravity dual, the gravity interpretation provides a natural preferred regularization scheme. With this scheme, for spherical entangling surfaces in $D=4$, the ratio is bounded as a consequence of the conformal collider bounds of [26,27]. Restricting to theories with holographic duals without higher derivative terms, the ratio turns out to be exactly equal to one.

Finally, in Sec. VI, we study how the capacity of entanglement behaves when conformal invariance is broken slightly. As a warm-up example, in the anisotropic Heisenberg XY spin chain, the equality of capacity and entropy emerges at critical domains. Moving away from criticality, the capacity of entanglement and the entanglement entropy are no longer equal and develop different subleading divergence structures. We then perform a more general analysis perturbing CFTs by relevant operators, following the strategy in [28–30].

We end with some concluding remarks.

II. DEFINITIONS AND RELATIONS

In this section, we introduce the concepts that will be the focus of later sections. As in thermodynamics, there will be

two different routes to defining these concepts. A “microscopic” route is to follow the standard definitions of quantum information theory. The second, “thermodynamic” route, follows from associating a Boltzmann distribution to the (reduced) density matrix and then formally applying thermodynamic definitions to quantum entanglement. We will begin with the first route, then follow the second route, and finally establish that the two different routes lead to the same concepts. Apart from some refinements, this section is mostly a review.

A. Concepts of quantum information theory

We start with a quantum system prepared into a pure or mixed state described by the density matrix ρ . Since we will mostly be interested in bipartite entanglement between a subsystem A and its complement, we assume that the density matrix is a reduced density matrix in the subsystem A . However, this is just for terminological convenience, most of the concepts obviously apply more generally.

The most familiar concept quantifying the purity of the state ρ , or the entanglement between the subsystem and its complement, is the entanglement entropy

$$S_{EE} = -\text{Tr}(\rho \log \rho). \quad (2.1)$$

Introducing the modular Hamiltonian $K = -\log \rho$, the entanglement entropy becomes the expectation value of K ,

$$S_{EE} = \text{Tr}(K\rho) = \langle K \rangle. \quad (2.2)$$

Another much studied measure of entanglement is given by the Rényi entropies S_α ,

$$S_\alpha = \frac{1}{1-\alpha} \log \text{Tr}(\rho^\alpha). \quad (2.3)$$

Recent work [7] suggested that in addition to the Rényi entropies S_α , it is interesting to study a modification, called “modular entropy” in [8],

$$\tilde{S}_\alpha \equiv \alpha^2 \partial_\alpha \left(\frac{\alpha-1}{\alpha} S_\alpha \right) = -\alpha^2 \partial_\alpha \left(\frac{1}{\alpha} \ln \text{Tr} \rho^\alpha \right). \quad (2.4)$$

The motivation in [7] comes from entanglement entropy in CFTs with a holographic gravitational dual, where the Ryu-Takayanagi formula relates the entanglement entropy to a volume of a dual minimal surface. In [7] it was found that the modular entropy also satisfies a holographic area law, with the interpretation¹

$$\tilde{S}_\alpha = \frac{\text{Area}(\text{Cosmic Brane})_\alpha}{4G_N}, \quad (2.5)$$

¹With α being a natural number.

where the “cosmic brane” is a domain wall with tension $(\alpha - 1)$ backreacting to the bulk geometry. In the limit $\alpha \rightarrow 1$, (2.5) reduces to the Ryu-Takayanagi formula.

In an earlier work [4] (see also [5]), (2.4) was used as a starting point to define a new quantity, capacity of entanglement C_E , to extend the thermodynamic relations found in quantum entanglement. In this paper we prefer to define it first in a microscopic, quantum information theoretic way, as a property of the state ρ ,

$$C_E(\rho) \equiv \text{Tr}(\rho(-\log \rho)^2) - (-\text{Tr}(\rho \log \rho))^2. \quad (2.6)$$

From this definition, written in terms of the modular Hamiltonian K , the capacity of entanglement is equal to the variance (the second cumulant) of K :

$$C_E(\rho) = \langle K^2 \rangle - \langle K \rangle^2. \quad (2.7)$$

The thermodynamic definition and its equality with (2.6) will be discussed in Secs. II B and II C.

A complete set of data associated with ρ is the entanglement spectrum [9], the complete set of eigenvalues $\{\lambda_m = e^{-\varepsilon_m}\}$ of the (reduced) density matrix ρ or the eigenvalues ε_m of the modular Hamiltonian K , along with their multiplicities g_m . There are other ways to encode the data of the spectrum $\{g_m, \varepsilon_m\}$. One alternative is to consider its moments or cumulants. For that, one can define a generating function $k(\alpha)$ as the analytic continuation² of the moments of the density matrix ρ ,

$$k(\alpha) = \text{Tr}(\rho^\alpha) = \sum_m g_m e^{-\alpha \varepsilon_m}. \quad (2.8)$$

By the usual rules, $k(\alpha)$ is a generating function for all the moments of the modular Hamiltonian:

$$\langle K^n \rangle = (-1)^n \left. \frac{d^n k(\alpha)}{d\alpha^n} \right|_{\alpha=1} = \langle (-\ln \rho)^n \rangle = \sum_m g_m e^{-\varepsilon_m} \varepsilon_m^n. \quad (2.9)$$

Formally, one may also refer to these as the moments of entanglement entropy,³ $S_{EE}^n = \langle K^n \rangle$. The partition function is related to the (analytically continued) Rényi entropies.

²In general, the analytic continuation may not be well defined for all α , but we will need the continuation only into a small neighborhood of $\alpha = 1$.

³In classical information theory, Shannon information $H = -\sum_i p(x_i) \log_2 p(x_i)$ quantifies the average information in messages composed of letters x_i with probabilities $p(x_i)$, generated at a source. In this context, the higher moments of H have a natural interpretation, characterizing the fluctuations in information. (A related concept is that of information loss in the evolution of chaotic dynamic systems, where the cumulants of H [31], and their generating function, the Rényi information [32] can be used to characterize dynamical chaos.)

Conversely, the analytic form of the Rényi entropies S_α determines $k(\alpha) = \exp[(1 - \alpha)S_\alpha]$, from which the eigenvalue spectrum can be extracted by a suitable expansion or integral transformation (see e.g., [33]).

The logarithm of the generating function,

$$\tilde{k}(\alpha) = \log k(\alpha) = (1 - \alpha)S_\alpha, \quad (2.10)$$

is the generating function for the cumulants (connected correlators),

$$\langle K^n \rangle_c = (-1)^n \left. \frac{d^n \tilde{k}(\alpha)}{d\alpha^n} \right|_{\alpha=1}. \quad (2.11)$$

If the explicit form of the Rényi entropies S_α is known, through (2.11) by applying derivatives, we can then easily calculate the capacity of entanglement C_E , which we defined in (2.6),

$$C_E = \langle K^2 \rangle_c = \left. \frac{d^2[(1 - \alpha)S_\alpha]}{d\alpha^2} \right|_{\alpha=1}. \quad (2.12)$$

The first derivative gives the first cumulant, which is S_{EE} . To summarize, there are four alternative ways to represent the entanglement data: the entanglement spectrum, the Rényi entropies, the moments $\langle K^n \rangle$ or the cumulants $\langle K^n \rangle_c$. It is possible to move from one description to the other. For example, knowing all the moments allows to construct the generating function $k(\alpha)$, which determines the entanglement spectrum. The different alternatives may be practical for highlighting different features of the entanglement data. The rest of the paper will focus on studying what features are captured by the second cumulant, the capacity of entanglement.

For pure states obviously all moments and cumulants vanish. For maximally mixed states,

$$\langle K^n \rangle = (\ln d)^n; \quad \langle K^{n>1} \rangle_c = 0, \quad (2.13)$$

in a Hilbert space with dimension $d < \infty$. The spectrum has only one d -fold degenerate eigenvalue $\varepsilon = \ln d$.

B. Thermodynamical definitions

In the previous section, we gave a quantum information theoretic definition and interpretation of the capacity of entanglement. In this section, we follow the original definition, by following the analogues of thermodynamical relations applied to quantum entanglement. We begin with the state $\rho = e^{-K}$ and introduce an inverse temperature β and compute a partition function

$$Z_\rho(\beta) = \text{Tr}(\rho^\beta) = \text{Tr}(e^{-\beta K}). \quad (2.14)$$

This is of course the same as the generating function $k(\alpha)$, with $\beta = \alpha$. The only reason for the two different notations

is to make a distinction between the thermodynamic and quantum information theoretic interpretations. $Z(\beta)$ is a more natural notation in the thermodynamical context, while $k(\alpha)$ refers to the quantum information theoretic context, hopefully this will not create confusion.

One can then proceed to introduce a free energy

$$F_\rho(\beta) = -\frac{1}{\beta} \log Z_\rho(\beta), \quad (2.15)$$

and an “internal” energy

$$E_\rho(\rho) = \frac{\partial}{\partial \beta} (\beta F_\rho(\beta)) = \frac{\text{Tr}(K e^{-\beta K})}{\text{Tr}(e^{-\beta K})} \equiv \langle K \rangle_\beta \quad (2.16)$$

which, as seen in the last line, is just the “thermal” expectation value⁴ of the modular Hamiltonian.

The thermodynamic definition of entropy satisfies the usual relation with the internal and free energy,

$$\begin{aligned} S_\rho(\beta) &= \beta^2 \frac{\partial F_\rho(\beta)}{\partial \beta} = \beta \langle K \rangle_\beta + \log Z_\rho(\beta) \\ &= \beta [E_\rho(\beta) - F_\rho(\beta)]. \end{aligned} \quad (2.17)$$

Finally, one defines a heat capacity,

$$C_\rho(\beta) = -\beta^2 \frac{\partial E_\rho(\beta)}{\partial \beta} = \beta^2 [\langle K^2 \rangle_\beta - \langle K \rangle_\beta^2]. \quad (2.18)$$

Except for the partition function, it is not manifestly clear how these thermodynamic definitions are related to the quantum information theoretic definitions of Sec. II A. That is the topic of the next section.

C. Relations between the thermodynamic and quantum information theoretic definitions

To begin with, from the density matrix $\rho = e^{-K}$ we construct a one-parameter family ρ_α of density matrices with proper normalization,⁵

$$\rho_\alpha = \frac{e^{-\alpha K}}{\text{Tr}(e^{-\alpha K})} = \frac{e^{-\beta K}}{Z_\rho(\beta)} \Big|_{\beta=\alpha}. \quad (2.19)$$

The key idea here will be that when we apply the quantum information theoretic definitions of Sec. II A (which were given for arbitrary states ρ) to the one-parameter family of states (2.19), we make contact with the thermodynamic definitions with $\beta = \alpha$.

We consider first the entanglement entropy. We observe that

⁴Notation: $\langle (\cdot) \rangle_\beta \equiv \frac{\text{Tr}((\cdot) e^{-\beta K})}{\text{Tr}(e^{-\beta K})}$.

⁵In [6], they were called escort density matrices, following the classical terminology of escort probability distributions.

$$\begin{aligned} S_{EE}(\rho_\alpha) &= -\text{Tr}[\rho_\alpha \log \rho_\alpha] = [\beta \langle K \rangle_\beta + \log Z_\rho(\beta)]|_{\beta=\alpha} \\ &= S_\rho(\beta)|_{\beta=\alpha}, \end{aligned} \quad (2.20)$$

verifying the relation between the two definitions of entropy.

On the other hand, we may consider the modular entropies \tilde{S}_α of the original state ρ . This gives another relation,

$$\begin{aligned} \tilde{S}_\alpha(\rho) &= -\alpha^2 \partial_\alpha \left(\frac{1}{\alpha} \log \text{Tr} \rho^\alpha \right) = [\beta \langle K \rangle_\beta + \log Z_\rho(\beta)]|_{\beta=\alpha} \\ &= S_\rho(\beta)|_{\beta=\alpha}, \end{aligned} \quad (2.21)$$

establishing $\tilde{S}_\alpha(\rho) = S_{EE}(\rho_\alpha) = S_\rho(\beta)|_{\beta=\alpha}$.

Finally, we show that the quantum information theoretic definition of the capacity of entanglement C_E (which was a property of an arbitrary state ρ , characterizing the variance of its spectrum), applied to the one-parameter family ρ_α , matches with the thermodynamic definition:

$$\begin{aligned} C_E(\rho_\alpha) &= \text{Tr}[\rho_\alpha (-\log \rho_\alpha)^2] - [-\text{Tr}(\rho_\alpha \log \rho_\alpha)]^2 \\ &= \{\beta^2 [\langle K^2 \rangle_\beta - \langle K \rangle_\beta^2]\}_{\beta=\alpha} = C_\rho(\beta)|_{\beta=\alpha}. \end{aligned} \quad (2.22)$$

This also establishes (setting $\alpha = 1$) the quantum information theoretic counterpart of the thermodynamic response-fluctuation relation between heat capacity and variance of entropy: for a state ρ the entanglement heat capacity and the variance of entanglement entropy satisfy

$$C_E(\rho) = \langle K^2 \rangle_c = \Delta S_{EE}^2(\rho), \quad (2.23)$$

which we used as the “microscopic” definition of C_E in (2.6).

Finally, later in the paper we will consider grand canonical ensembles of either identical bosonic or fermionic particles, and there it will be convenient to use the expression of the capacity or variance in terms of the expectation values of the occupation numbers n_l as a single sum,

$$C_E = \Delta S_{EE}^2 = \sum_l (\log(1 \pm n_l) - \log(n_l))^2 (1 \pm n_l) n_l, \quad (2.24)$$

where we choose the + signs for bosonic and – signs for fermionic systems.

D. Variance, fidelity susceptibility and quantum Fisher information

The capacity or variance (2.6), (2.12), can also be related to other concepts of quantum information: fidelity susceptibility (FS) and quantum Fisher information (QFI). These concepts are important in quantum metrology, and can be used in various contexts, such as Bose-Einstein condensation, temperature estimates, and quantum phase

transitions, see e.g., [34,35], and the recent extensive review [36] (and references therein). Here we give only a brief review for the purpose of making a comparison to the capacity of entanglement C_E .

Consider deformations of a density matrix, parametrized by a continuous parameter θ . This essentially gives a one-parameter flow $\rho(\theta)$ in the state space. One measure of a “distance” between two density matrices ρ, σ is the (Uhlmann) fidelity $F(\rho, \sigma)$ [37],

$$F(\rho, \sigma) = \text{Tr} \sqrt{\sqrt{\rho} \sigma \sqrt{\rho}}. \quad (2.25)$$

Then $F(\rho(\theta), \rho(\theta + \epsilon))$ can be used as a measure of how much the density matrix changes along an infinitesimal flow $\theta + \epsilon$. Since the fidelity is at maximum value 1 at $\epsilon = 0$, its Taylor series expansion is

$$F(\rho(\theta), \rho(\theta + \epsilon)) = 1 - \frac{1}{2} \chi_F(\theta) \epsilon^2 + \dots \quad (2.26)$$

where the coefficient

$$\chi_F(\theta) \equiv - \frac{\partial^2 F(\rho(\theta), \rho(\theta + \epsilon))}{\partial \epsilon^2} \Big|_{\epsilon=0} \quad (2.27)$$

is called the fidelity susceptibility. For example, consider θ as a control parameter driving the ground state of a system across a quantum phase transition. Then the fidelity susceptibility is an interesting observable, manifesting singular behavior that can be used to characterize the quantum phase transition.

A slightly more general approach is to compare ρ and $\rho + \epsilon \delta \rho$, and work in a basis where ρ is diagonal with eigenvalues λ_i . Note however that in this basis the perturbation $\delta \rho$ need not be diagonal. A brute force computation gives then

$$F(\rho, \rho + \delta \rho) = 1 - \frac{\epsilon^2}{4} \sum_{ik} (\delta \rho)_{ik} (\delta \rho)_{ki} \frac{1}{\lambda_i + \lambda_k} + \mathcal{O}(\epsilon^3) \quad (2.28)$$

from which we read the fidelity susceptibility

$$\chi = \frac{1}{2} \sum_{ik} (\delta \rho)_{ik} (\delta \rho)_{ki} \frac{1}{\lambda_i + \lambda_k}. \quad (2.29)$$

We will shortly connect this to capacity of entanglement but before doing that we first discuss another related concept, quantum Fisher information. Let us specify a point $\theta = 0$ and consider small deformations about it. In quantum measurements, one is often interested in detecting the parameter θ by finding an observable $\hat{\theta}$, a locally unbiased estimator with the normalization

$$\frac{\partial}{\partial \theta} \text{Tr}(\rho(\theta) \hat{\theta})|_{\theta=0} = 1. \quad (2.30)$$

The accuracy of the estimates of θ is characterized by the variance $\Delta \theta^2$, which is bounded from below by the quantum version of the Cramér-Rao bound [38],

$$\Delta \theta^2 \geq \frac{1}{N g(\theta)} \quad (2.31)$$

where N is the number of samples and the lower bound is characterized by the quantum Fisher information $g(\theta)$. To define it, consider first the so-called symmetric logarithmic derivative $L(\theta)$ of $\rho(\theta)$, defined by

$$\frac{\partial \rho(\theta)}{\partial \theta} = \frac{1}{2} (L \rho + \rho L). \quad (2.32)$$

The quantum Fisher information $g(\theta)$ is then defined as

$$g(\theta) = \text{Tr}(\rho(\theta) L^2(\theta)) = \text{Tr} \left(\frac{\partial \rho(\theta)}{\partial \theta} L(\theta) \right). \quad (2.33)$$

For an explicit formula, consider again a small deformation $\rho + \epsilon \delta \rho$, in a basis where ρ is diagonal (but $\delta \rho$ need not be). Then solve the equation

$$\delta \rho = \frac{1}{2} (L \rho + \rho L) \quad (2.34)$$

for L , in terms of the eigenvalues λ_i of ρ we get

$$L_{ij} = \frac{2}{\lambda_i + \lambda_j} (\delta \rho)_{ij}. \quad (2.35)$$

It is now straightforward to establish an equality relating the Fisher information to the fidelity susceptibility (2.29),

$$g = \text{Tr}(\rho L^2) = 4\chi. \quad (2.36)$$

Coming back to the variance $\langle K^2 \rangle_c$, let us consider a particular deformation of the density matrix $\rho = e^{-K}$ by deforming in the direction of (imaginary) modular flow,

$$\rho(\theta) = e^{-(1+\theta)K} / (\text{Tr} e^{-(1+\theta)K}), \quad (2.37)$$

in other words the family of escort matrices (2.19). Computing the fidelity $F(\rho(0), \rho(\theta))$ to second order in θ (the normalization factor must also be expanded), and setting $\theta = 0$, yields

$$C_E = \langle K^2 \rangle_c = 4\chi_F(0) = g, \quad (2.38)$$

i.e., the capacity C_E is equal to the fidelity susceptibility for this particular flow and the quantum Fisher information for estimating the parameter θ . Alternatively, expanding

$\rho(\theta) \approx \rho(0) + \theta \delta \rho + \mathcal{O}(\theta^2)$, if $\rho(0)$ is diagonal with eigenvalues λ_i , then $\delta \rho$ is diagonal with eigenvalues $\lambda_i \log \lambda_i - \lambda_i \sum_j \lambda_j \log \lambda_j$, and then (2.29), (2.36) yield (2.38).

The equality (2.38) is well-known, see e.g., [17]. In a finite temperature system, when the density matrix is that of a canonical ensemble, the above quantities are called the thermal fidelity susceptibility *etc.*, and also equal to the heat capacity [39]. However, we emphasize that the first equality in (2.38) assumes the particular flow (2.37). For example, suppose that we deform the modular Hamiltonian by a perturbation V , to consider a flow

$$\tilde{\rho}(\theta) = e^{-K - \theta V} / (\text{Tr} e^{-K - \theta V}). \quad (2.39)$$

This leads to a more complicated expansion of the fidelity, since in general V , K do not commute, and the second equality in (2.38) does not hold.

E. Relation to thermal heat capacity

So far we have mostly been keeping the discussion at a general level, without specifying the origin of the quantum state ρ . In the remainder of the paper, we will narrow our focus to bipartite entanglement in quantum systems, where the system is first prepared to a generic quantum state ρ_0 , then partitioned into a subsystem A entangled with its complement B , and then study the reduced density matrix $\rho_A = \text{Tr}_B \rho_0$. The entanglement entropy is the von Neumann entropy of ρ_A . In this section, we consider specifically the entanglement entropy for CFT initial vacuum states and spherical entangling surfaces, in the context of [40,41]. In these papers, the authors discuss the interpretation of EE and Rényi entropies, by mapping the causal development of the enclosed ball A to a hyperbolic cylinder $\mathbb{R} \times H^{d-1}$, where the CFT in the vacuum in the original spacetime becomes a CFT in H^{d-1} in a thermal bath, with temperature

$$T_0 = \frac{1}{2\pi R}, \quad (2.40)$$

where R is the radius of the original entangling sphere and the curvature scale of H^{d-1} . The entanglement entropy is then mapped to standard thermal entropy $S_{\text{therm}}(T_0)$ of the CFT in the thermal bath, similarly the Rényi entropies are mapped to those of the thermal ensemble.

We consider the capacity of entanglement C_E of the ball-like subsystem. We show that it maps to the standard thermal heat capacity C_{therm} of the thermal CFT. We start from⁶ (2.12), but do not yet take the limit $\alpha \rightarrow 1$ but consider

$$C(\alpha) = \alpha^2 \partial_\alpha^2 \ln \text{Tr} \rho_A^\alpha = \alpha^2 \partial_\alpha^2 [(1 - \alpha) S_\alpha], \quad (2.41)$$

where S_α is the Rényi entropy. On the other hand, it is related to the free energy F on the hyperbolic sphere by [41]

⁶We use the notation α where [41] uses q .

$$S_\alpha = \frac{\alpha}{1 - \alpha} \frac{1}{T_0} [F(T_0) - F(T_0/\alpha)], \quad (2.42)$$

so

$$C(\alpha) = \alpha^2 \partial_\alpha^2 \left[-\frac{\alpha}{T_0} F\left(\frac{T_0}{\alpha}\right) \right]. \quad (2.43)$$

Writing $T_\alpha \equiv T_0/\alpha$ and $\beta_\alpha = 1/T_\alpha$, we get

$$C(\alpha) = -\beta_\alpha^2 \frac{\partial}{\partial \beta_\alpha} E(T_\alpha) = \frac{\partial E(T_\alpha)}{\partial T_\alpha}. \quad (2.44)$$

Thus,

$$C_E = \lim_{\alpha \rightarrow 1} C(\alpha) = \frac{\partial E(T_0)}{\partial T_0} = C_{\text{therm}}, \quad (2.45)$$

i.e., the capacity of entanglement C_E of ρ_A becomes the usual heat capacity C_{therm} of the thermal CFT on the hyperbolic sphere, as was expected.⁷ The temperature is parametrized by the radius of the entangling sphere, (2.40). This also guarantees that the holographic dual interpretation with a topological black hole in the bulk will be the heat capacity (associated with the horizon) of the black hole. Then, dialing the temperature of the black hole translates back in the original spacetime to the imaginary modular flow among the escort matrices accompanying the reduced density matrix ρ_A of the subsystem A .

It should be noted that in Appendix A, [41] considered various inequalities satisfied by the Rényi entropy, and interpreted one of them [their (A4)] as a result of proportionality to a heat capacity which is positive. These inequalities were studied and proven in [6].

Interim summary. We have established a string of equalities between the variance of the entanglement spectrum $\langle K^2 \rangle_c$, the variance of entanglement entropy, the capacity of entanglement, a reduced fidelity susceptibility and quantum Fisher information,

$$\langle K^2 \rangle_c = \Delta S_{EE}^2 = C_E^* = 4\chi_F = g. \quad (2.46)$$

However, the third equality (marked with an asterisk) holds for a particular (modular) flow among density matrices. Hereafter, we will mostly adopt the notation C_E , and we will be interested in comparisons to the entanglement entropy S_{EE} , in a given state ρ .

⁷A precursor of such a connection was discussed in [42], which introduced a heat capacity of entanglement across a planar entangling surface, relating it by a coordinate transformation to the thermal heat capacity of Rindler radiation, also finding a linear area scaling.

III. PROPERTIES OF CAPACITY OF ENTANGLEMENT

We have introduced the cumulants of the modular Hamiltonian, and interpreted the second cumulant as the capacity of entanglement or the variance of entanglement entropy. In this section we study some properties of these concepts.

A. An upper bound on the capacity of entanglement

In a system with a finite Hilbert space of dimension N , the von Neumann entropy takes values between 0 and $S_{\max} = \ln N$. It is interesting to derive similar bounds for the higher cumulants of the spectrum of ρ . In this section we derive an upper bound for the capacity of entanglement $C_E = \langle K^2 \rangle_c$.

We may assume that we have diagonalized the density matrix. We denote the eigenvalues λ_i , with the degenerate eigenvalues also carrying separate labels. There are two types of questions to consider: (i) for a given entropy S , what is the maximum variance $\langle K^2 \rangle_c$, or (ii) find the maximum of $\langle K^2 \rangle_c$ with no restriction on the entropy.

We start with the problem (i). It leads to the variational problem to extremize the functional

$$F = \sum_i \lambda_i (\log \lambda_i)^2 + A \left(S - \sum_i \lambda_i \log \lambda_i \right) + \sum B_i \theta(-\lambda_i) + C \left(\sum_i \lambda_i - 1 \right), \quad (3.1)$$

where the Lagrange multipliers A , B_i and C enforce that the entropy takes a given value, that the λ_i are non-negative, and that the sum of the eigenvalues is equal to one, respectively. Note that F is not equal to $\langle K^2 \rangle_c$, instead $F = \langle K^2 \rangle_c + S^2$, but since S is fixed extremizing F is the same as extremizing $\langle K^2 \rangle_c$.

Differentiating F with respect to λ_i yields

$$\log^2 \lambda_i + (2 - A) \log \lambda_i - A - B_i \delta(\lambda_i) + C = 0. \quad (3.2)$$

We see that if $\lambda_i \neq 0$, this is a quadratic equation for $\log \lambda_i$ with at most two solutions for λ_i . In addition, some of the λ_i could be zero, for which this equation is somewhat ill-defined, but can formally be solved with $B_i = 0$ and C infinite.

We now first move on to the alternative problem (ii), which turns out to lead to a simple result, before turning back to (i). If we do not fix S , we need to maximize

$$\tilde{F} = \sum_i \lambda_i (\log \lambda_i)^2 - \left(\sum_i \lambda_i \log \lambda_i \right)^2 + \sum B_i \theta(-\lambda_i) + C \left(\sum_i \lambda_i - 1 \right). \quad (3.3)$$

Differentiating with respect to λ_i now leads to

$$\log^2 \lambda_i + 2 \log \lambda_i - 2(1 + \log \lambda_i) \left(\sum_i \lambda_i \log \lambda_i \right) + B_i \delta(\lambda_i) + C = 0. \quad (3.4)$$

Also in this case, having three different nonzero values among the λ_i would lead to a contradiction, as all three would be solutions of a quadratic equation for $\log \lambda_i$ with the same coefficients.

We therefore conclude that for either optimization problem, the nonzero eigenvalues can take at most two different values. Since a zero eigenvalue does not contribute to either entanglement entropy or the variance, we will ignore the zero eigenvalues for now.

We will denote the two non-zero eigenvalues by λ_1 and λ_2 . If we call the corresponding occupation numbers n_1 and n_2 , then $n_1 \lambda_1 + n_2 \lambda_2 = 1$. We can express n_1 and n_2 in terms of λ_1 and λ_2 and when we do this we find

$$C_E = \langle K^2 \rangle_c = -(S + \log \lambda_1)(S + \log \lambda_2), \quad (3.5)$$

which shows that in order to maximize C_E , we should make λ_1 and λ_2 as different as possible.

It will be convenient to introduce the variables

$$x = n_1 \lambda_1, \quad y = \frac{n_1}{n_2} \quad (3.6)$$

so that we can express all quantities in terms of x , y and $N = n_1 + n_2$. We find that

$$S = \log N - x \log \frac{x}{y(1-x)} - \log(1-x)(1+y) \quad (3.7)$$

and

$$C_E = \langle K^2 \rangle_c = x(1-x) \left[\log \left(y \frac{1-x}{x} \right) \right]^2. \quad (3.8)$$

Let us try to maximize $\langle K^2 \rangle_c$. We notice that first of all we want to minimize y . The minimum value would be $y = 0$, but this is just the maximally entangled state, for which the capacity of entanglement vanishes. Instead, we set $y = 1/(N-1)$. Then, the maximum of $\langle K^2 \rangle_c$ is found at x_N , which is the solution of

$$\log \left(\frac{1-x}{x} \right) - \frac{2}{1-2x} - \log(N-1) = 0. \quad (3.9)$$

We then find an upper bound for the variance,

$$\langle K^2 \rangle_c \leq \frac{4x_N(1-x_N)}{(1-2x_N)^2}. \quad (3.10)$$

Solving the transcendental equation (3.9) iteratively, we obtain for large N

$$x_N \approx \frac{1}{2} + \frac{1}{\log(N-1)} + \dots \quad (3.11)$$

and

$$C_{E,\max} = \langle K^2 \rangle_{c,\max} \approx \frac{1}{4} \log^2 N + 1 + \dots \approx \frac{1}{4} S_{\max}^2. \quad (3.12)$$

Notice that we did not keep S fixed when we varied x and y . However, for the above values of x and y the entropy for large N remains of the order $S \sim \log N$ so we expect that it remains true that $C_{E,\max}$ is proportional to S_{\max}^2 (with a different numerical constant) if we were to vary $\langle K^2 \rangle_c$ keeping S fixed and of the order $\log N$. It would be interesting to know if states with maximum capacity of entanglement have special properties, but we have no specific insights at this point.

Formally, the inequality

$$C_E = \langle K^2 \rangle_c = \Delta S_{EE}^2 \leq \frac{1}{4} (S_{\max})^2 \quad (3.13)$$

resembles Popoviciu's inequality on variances [43], which states that the variance of a random variable X with a range from $\inf(X) = m$ to $\sup(X) = M$ satisfies

$$\Delta X^2 \equiv \text{var}(X) \leq \frac{1}{4} (M - m)^2. \quad (3.14)$$

If we were to interpret the entropy S_{EE} as a random variable with the upper bound S_{\max} and the lower bound 0, the inequality (3.13) would indeed superficially look like a special case of Popoviciu's inequality. However, the spectrum of the modular Hamiltonian depends on the probability distribution, and has no finite upper bound. To clarify, we recast our result in the language of classical information theory. Let X be a random variable with N outcomes $\{x_1, \dots, x_N\}$, with a discrete probability distribution $p(x_i)$. Then move to a new random variable, the information content $I(X) = -\log_2 p(X)$. The expectation value is the Shannon information, $H = E(I(X))$, with $0 \leq H \leq H_{\max} = \log_2 N$. The same variational calculation as we performed above, changing the notation, gives an estimate for the variance of $I(X)$

$$\Delta I(X)^2 \equiv \text{var}(I(X)) \leq \frac{1}{4} (H_{\max})^2. \quad (3.15)$$

This is not a special case of Popoviciu's inequality: unlike for X , for the random variable $I(X)$ the supremum depends on the probability distribution $p(x)$, with $\sup I(X) = \infty$. Instead, what appears in (3.15) is the supremum of the expectation value. We are unaware of whether (3.15) is a new result in classical information theory, or previously known.

Note also that the upper bound applies for a finite system. In the thermodynamic limit, with the system size becoming infinite, it would be interesting to study if the capacity of entanglement displays interesting nonanalyticity e.g., when the system is undergoing a quantum phase transition.

B. Bipartite particle fluctuations and entanglement fluctuations

In this section, we discuss scenarios where it is natural to consider fluctuations in entanglement entropy and its cumulants. We consider entanglement in a bipartitioned system. There has been growing interest in developing strategies to measure entanglement. An interesting proposal relates entanglement in a bipartitioned system (between the subsystem A and its complement) to fluctuations of a conserved $U(1)$ charge, such as the particle number, across the partition boundary [19–23], for systems that can be mapped to noninteracting fermions. The particle number in the subsystem A , N_A , thus becomes a random variable, and so does the entropy as entangled charges wander in and out. There are detailed results relating the fluctuations in the particle number N_A in the subregion A to the entanglement entropy S_{EE} and even to the Rényi entropies S_α . The probability distribution of the random number N_A is characterized by its cumulants n_m , defined by

$$n_m = (-i\partial_\lambda)^m \log \chi(\lambda)|_{\lambda=0}, \quad (3.16)$$

with the generating function

$$\chi(\lambda) = \langle e^{i\lambda N_A} \rangle, \quad (3.17)$$

where the expectation value is computed with respect to the initial state of the full system. Thus, for example the second cumulant n_2 (also called the fluctuations) is

$$n_2 = \langle N_A^2 \rangle - \langle N_A \rangle^2. \quad (3.18)$$

Given the relation between entanglement and particle number fluctuations, we expect the probability distributions of these two random variables to be related, and in particular to be able to express all the cumulants of the entanglement entropy (the modular Hamiltonian) in terms of the cumulants of the particle number. Let us show this in more detail. We emphasize that this only applies to systems that can be mapped to noninteracting fermions. In this context, the entanglement Rényi entropies have the following series expansion in terms of the n_m :

$$S_\alpha(A) = \sum_{k=1}^{\infty} s_k^{(\alpha)} n_{2k} \quad (3.19)$$

with the coefficients

$$s_k^{(\alpha)} = \frac{(-1)^k (2\pi)^{2k} 2\zeta(-2k, (1+\alpha)/2)}{(\alpha-1)\alpha^{2k}(2k)!}, \quad (3.20)$$

where $\zeta(s, a)$ is the Hurwitz zeta function. Note that only the even cumulants n_{2k} appear in the series. In principle, the series is valid for real values of $\alpha > 1$, but one has to be careful about its convergence. In [44], it was shown that, for systems equivalent to noninteracting fermion gases, the series (3.19) is well behaved, because the first cumulant n_2 contributes the leading $N^{(d-1)/d} \log N$ asymptotic behavior, while the higher cumulants $n_{k>2}$ contribute only subleading behavior, N being the total particle number. Thus the series essentially truncates.

In this case, we can view $\tilde{k}(\alpha) = (1-\alpha)S_\alpha(A)$ as a well behaved generating function for the cumulants of the entanglement entropy/modular Hamiltonian. In particular, the entanglement entropy and the capacity of entanglement have the expansions

$$\begin{aligned} S_{EE}(A) &= \sum_{k=1}^{\infty} \frac{(2\pi)^{2k} |B_{2k}|}{(2k)!} n_{2k} \\ &= \frac{\pi^2}{3} n_2 + \frac{\pi^4}{45} n_4 + \frac{2\pi^6}{945} n_6 + \dots, \\ C_E(A) &= \frac{\pi^2}{3} n_2 + \sum_{k=2}^{\infty} \frac{2(2\pi)^{2k} |B_{2k}|}{(2k-1)!} n_{2k} \\ &= \frac{\pi^2}{3} n_2 + \frac{8\pi^4}{45} n_4 + \frac{24\pi^6}{945} n_6 + \dots, \end{aligned} \quad (3.21)$$

in the above B_{2k} are Bernoulli numbers. From the results, we immediately see that if the particle fluctuations are Gaussian distributed ($n_{m>3} = 0$), the entanglement entropy is equal to its variance. Furthermore, in [44] it was noted that in the $N \rightarrow \infty$ limit the Rényi entropies remarkably satisfy

$$\frac{S_\alpha(A)}{n_2} = \frac{\pi^2}{6} (1 + \alpha^{-1}) + \dots, \quad (3.22)$$

where $+\dots$ indicates corrections that vanish. A consequence of this is that

$$C_E(A) = S_{EE}(A) \quad (3.23)$$

in the $N \rightarrow \infty$ limit. Note that the dimension d of the system was generic in the derivation. However, here the size of the subregion A is independent of N and is kept fixed.

As a diversion, let us make some very speculative but perhaps inspirational remarks regarding systems that have holographic gravity duals. In this setting, we might make the following argument for invoking a large N limit and Gaussian statistics. Holographic gauge/gravity duality is simplified in a limit where the bulk gravity system is classical, corresponding to a large \mathcal{N} limit in the dual theory at the AdS boundary, where \mathcal{N} is the number of

constituent degrees of freedom (d.o.f.). In the dual theory, one could then focus on the average number of d.o.f. fluctuating in the subregion A ,

$$N_A = \frac{1}{\mathcal{N}} \sum_{i=1}^{\mathcal{N}} N_A^{(i)}, \quad (3.24)$$

where $N_A^{(i)}$ is the particle number of the species i in the subregion A . Viewing each $N_A^{(i)}$ as an independent random variable, not necessarily having identical distributions, under suitable conditions (e.g., Lyapunov conditions), the central limit theorem applies and in the large \mathcal{N} limit the average particle number N_A becomes Gaussian distributed. If the quantum entanglement is attributed to the average particle number fluctuations about the subregion, from above one could deduce that $C_E = S_{EE}$. An additional complication is that in gauge/gravity duality, one typically works in the large 't Hooft coupling limit so that the bulk spacetime is weakly curved. The dual theory on the boundary is then strongly coupled, so the concept of “particles” is lost and the definition of a “particle number” becomes less clear. Another caveat is that the definition of entanglement for gauge invariant observables is more complicated [45,46].

In Sec. VI A, we study a system which can be mapped to noninteracting fermions. We consider ground state entanglement in a Heisenberg XY spin chain at different phases, with the ground state adjusting to the changes in the parameters. Its Rényi entropies were computed in [33,47], although they worked in a double scaling limit $N \rightarrow \infty$, $L \rightarrow \infty$ with L/N fixed, where L is the size of the subsystem. Therefore, the above large N result from particle number fluctuations is not directly applicable. In particular, we will find that the entanglement entropy is not always equal to the capacity of entanglement. This system will also work as an introduction to later sections where we study CFTs and their perturbations with relevant operators.

C. A simple example where capacity equals entropy

The large N limit in the previous section led to the equality $C_E = S_{EE}$. Later, we will arrive several times at the same result when we discuss entanglement in CFTs. It is interesting to ask what probability distributions could lead to such an equality. Here is one concrete example. It will be convenient to write $\lambda = e^{-E}$ and to use a density of states $\rho(E)$. For a finite dimensional density matrix with eigenvalues λ_i , the density of states would be $\rho(E) = \sum_i \delta(E + \log \lambda_i)$. Then,

$$\begin{aligned} \int dE \rho(E) e^{-E} &= 1, \\ \int dE E \rho(E) e^{-E} &= S_{EE}, \\ \int dE E^2 \rho(E) e^{-E} &= C_E - S_{EE}^2. \end{aligned} \quad (3.25)$$

If we take, for example, $\rho(E) = E^k/k!$, we obtain $C_E = S_{EE} = k + 1$, which is indeed an example where $C_E = S_{EE}$.

D. Gravity dual of the capacity of entanglement

Results in [7] can be used to find the gravity dual of the capacity of entanglement. This was done in [6], but here we present a simplified derivation of the result, and point out some additional properties.

Starting from the modular entropy (2.4), we can rewrite the capacity of entanglement as a first derivative,

$$C_E = \Delta S_{EE}^2 = -\frac{d\tilde{S}_\alpha}{d\alpha}\bigg|_{\alpha=1}. \quad (3.26)$$

The modular entropy \tilde{S}_α is given by the area of a suitably backreacted cosmic brane as in (2.5) and to determine that we need to consider the action for a brane coupled to gravity

$$I = -\frac{1}{16\pi G_N} \int \sqrt{g} R + I_{\text{matter}} + \frac{\alpha-1}{4\alpha G_N} \int_{\text{brane}} \sqrt{h} \quad (3.27)$$

with h the induced metric on the brane. To leading order, the cosmic brane is just a minimal surface. To first subleading order, we need to take backreaction from the tension of the brane into account. The brane yields a source for the metric, and to first order in the source this affects the metric everywhere in spacetime through the bulk graviton propagator. It could also affect the matter fields to first order, in case the kinetic terms for the graviton and matter fields mix in this background.

At higher order, we would need to include more complicated Witten diagrams to understand the backreaction. If we only consider linear metric fluctuations, the location of the minimal surface does not change, only its area changes. This is because the minimum area surface is an extremum of the area functional and therefore its location does not change under first order perturbations.

The first-order variation of the area is given by

$$\delta \frac{\text{Area}}{4G_N} = \frac{1}{4G_N} \int \sqrt{h} \frac{1}{2} h^{ij} \delta h_{ij}. \quad (3.28)$$

The field equation of (3.27) is

$$-\frac{1}{16\pi G_N} \left(R_{\mu\nu} - \frac{1}{2} R g_{\mu\nu} \right) + \frac{1}{2} T_{\mu\nu}^{\text{matter}} - \frac{\alpha-1}{8\alpha G_N} h_{ij} \delta(\text{brane}) = 0, \quad (3.29)$$

where a simplified notation has been used. A more precise way to write the last term, using the explicit expression for the induced metric h_{ij} in terms of the embedding $x^\mu(\sigma^i)$ of the surface,

$$h_{ij} = g_{\mu\nu}(x^\mu(\sigma^i)) \frac{\partial x^\mu}{\partial \sigma^i} \frac{\partial x^\nu}{\partial \sigma^j}, \quad (3.30)$$

reads

$$-\frac{\alpha-1}{8\alpha G_N} \int d^{d-1} \sigma_i \frac{\sqrt{h}}{\sqrt{g}} h^{ij} g_{\mu\rho} \frac{\partial x^\rho}{\partial \sigma^i} \frac{\partial x^\eta}{\partial \sigma^j} g_{\eta\nu} \delta(x^\mu - x^\mu(\sigma^i)). \quad (3.31)$$

In order to not clutter the notation, we will keep using the imprecise notation in (3.29) but write the final answer in a more accurate way.

We need to solve (3.29) to first order in $(\alpha-1)$. To leading order, the metric obeys the field equations in the bulk, and we need to linearize around the background in order to find the first order variation, $g_{\mu\nu} = g_{\mu\nu}^{\text{background}} + \delta g_{\mu\nu}$

$$-\frac{1}{16\pi G_N} \mathcal{D}_{\text{background}} \delta g_{\mu\nu} - \frac{\alpha-1}{8\pi G_N} g_{ij}^{\text{background}} \delta(\text{brane}) + \mathcal{O}((\alpha-1)^2) = 0 \quad (3.32)$$

with \mathcal{D} some second order operator representing the kinetic term of the graviton. Therefore, if G is the graviton propagator in the background, we get

$$\delta g_{\mu\nu}(x) = -\frac{1}{8G_N} \int dx' \sqrt{g(x')} G_{\mu\nu}^{\alpha\beta}(x, x') \times \frac{\alpha-1}{\alpha} g_{\alpha\beta}^{\text{background}}(x') \delta(\text{brane}). \quad (3.33)$$

Inserting this in the variation (3.28), we are left with a double integral over the minimal surface

$$C_E = \frac{1}{64G_N^2} \int dx \sqrt{g(x)} \int dx' \sqrt{g(x')} g^{ij} G_{ij}^{kl}(x, x') g_{kl}. \quad (3.34)$$

Notice that the indices of the graviton propagator are contracted with the metric on the brane, in other words we are effectively taking a trace, but only in the directions along the brane.

A more precise version of (3.34) in terms of the embedding $x^\mu(\sigma^i)$ is

$$C_E = \frac{1}{64G_N^2} \int d^{d-1} \sigma \sqrt{h(\sigma)} \int d^{d-1} \sigma' \sqrt{h(\sigma')} h^{ij} \times G_{ij;kl}(x(\sigma), x(\sigma')) h^{kl} \quad (3.35)$$

with

$$G_{ij;kl}(x(\sigma), x(\sigma')) = G_{\mu\nu;\rho\eta}(x(\sigma), x(\sigma')) \frac{\partial x^\mu}{\partial \sigma^i} \frac{\partial x^\nu}{\partial \sigma^j} \frac{\partial x^\rho}{\partial \sigma'^k} \frac{\partial x^\eta}{\partial \sigma'^l}. \quad (3.36)$$

This result was derived in [6], via a different route.

We note in passing that the result (3.35) looks very familiar when phrased in terms of the (reduced) fidelity susceptibility associated with the flow (2.37) as in Sec. II. Then,⁸

$$\chi_F = \int_{\Sigma} d\vec{x} \int_{\Sigma} d\vec{x}' g^{ij} g_{kl} G_{ij}^{kl}(\vec{x}, \vec{x}'), \quad (3.37)$$

a relationship between response coefficient and fluctuations resembling that of the magnetic susceptibility and order parameter fluctuations in Landau-Ginzburg theory,

$$\chi = \int d\vec{x} \int d\vec{x}' \delta^{ij} C_{ij}(\vec{x}, \vec{x}'), \quad (3.38)$$

with $C_{ij}(\vec{x}, \vec{x}') = \langle m_i(\vec{x}) m_j(\vec{x}') \rangle_c$. The analogue is satisfying, viewing bulk gravity as an effective theory for the quantum theory on the boundary.

One might have thought that one can simply decouple this piece of the graviton propagator by restricting to the conformal mode, but that would involve a trace over all indices, not just those along the brane. We have not attempted to compute (3.34) in an actual example but can relate this expression to CFT computations that have appeared previously in the literature [24,25] for the expansion of the Rényi entropy S_{α} around $\alpha = 1$ for a spherical entangling surface. We first notice that the first-order variation of entanglement entropy itself reads

$$\delta S_{EE} = \frac{1}{8G_N} \int d^{d-1} \sigma \sqrt{h} h^{ij} \delta h_{ij}, \quad (3.39)$$

and therefore (3.35) can be interpreted as a two-point function

$$C_E = \Delta S_{EE}^2 = \langle \delta S_{EE} \delta S_{EE} \rangle, \quad (3.40)$$

where we view δh_{ij} as a fluctuating bulk field. This clarifies more directly the relation between the bulk computation and the boundary computation in terms of correlation functions of the modular Hamiltonian K . Indeed, if we interpret the first law of entanglement entropy as an operator statement (see e.g., [48,49]), so that

$$\delta S_{EE} \equiv K, \quad (3.41)$$

then the equivalence between $\langle \delta S_{EE} \delta S_{EE} \rangle$ and $\langle KK \rangle$ becomes manifest. In this way, we can also much more directly derive bulk expressions for the higher moments of K , they simply become higher order bulk correlators of δS_{EE} , and in this way perturbatively prove the result of [7].

The above computations also nicely illustrate the relation of quantum entanglement to gravity: we can see that entanglement fluctuations are interpreted as self-gravitation, with the two-point function (3.40) as

the leading order integrated self-gravitation of the bulk surface⁹

E. Direct CFT computation

As we just described, a more direct CFT computation of the capacity of entanglement and higher moments of the entanglement spectrum involves the computation of n -point functions of the modular Hamiltonian. For spherical regions, the modular Hamiltonian has a local expression

$$K_A = 2\pi \int_{|x| < R} d^{d-1} x \frac{R^2 - r^2}{2R} T_{tt}(x) \quad (3.42)$$

and the capacity of entanglement then involves a double integral of the 2-point function $\langle T_{tt}(x) T_{tt}(x') \rangle$. In the case $d = 2$, using a slightly *ad hoc* way to treat the divergent integral,¹⁰

$$C_E = \langle K_A^2 \rangle_c = \int_{-R}^R dx \int_{-R}^R dx' \frac{(R^2 - x^2)(R^2 - x'^2)}{4R^2} \frac{c}{(x - x')^4} \quad (3.43)$$

$$= \int_{-R(1-\varepsilon)}^{R(1-\varepsilon)} dx' c \frac{R(R^2 - x'^2)}{3(R - x')^2(R + x')^2} \quad (3.44)$$

$$= \frac{2c}{3} \text{ArcTanh}(1 - \varepsilon) \quad (3.45)$$

$$= \frac{c}{3} \left(\log \frac{2}{\varepsilon} \right) + o(\varepsilon) \quad (3.46)$$

$$= \frac{c}{3} \log \frac{L}{a}. \quad (3.47)$$

In the above, we used the well-known $\langle T(z) T(0) \rangle \sim \frac{c}{2z^4}$ formula and also set a UV regulator $\varepsilon = \frac{2a}{L}$ by hand. In the first step, we did the x integral assuming the contour had been shifted off the real axis. The second integral is then performed with an explicit cutoff.

One can systematically study higher cumulants and see that they are given by the terms that appear in the higher order correlation functions of the energy momentum tensor. This is explored in quite some detail in [25], see also Sec. VB, in particular the relation between the second derivative of the Rényi and the coefficient that appear in the

⁹Given the equality (2.46) of the capacity of entanglement, fidelity susceptibility and quantum Fisher information under modular flow, the corresponding bulk interpretations of the three should also be related. However, on the face of it, proposals for the latter two look quite different [17,18,50,51]. It would be interesting (but beyond the scope of this work) to understand in more detail the relation of the bulk interpretations when (2.46) holds.

¹⁰It should be remembered that $T = 2\pi T_{zz}$ and $\bar{T} = 2\pi T_{\bar{z}\bar{z}}$.

⁸Absorbing the normalization factor into the graviton propagator.

three-point function of the energy momentum tensor is worked out in this paper. This paper also discusses a particular regularization of the integrated correlation functions of the type above, but only in a simple example, and proposes to use dimensional regularization whose connection to standard regulators is obscure.

IV. STATE DEPENDENCE OF THE CAPACITY OF ENTANGLEMENT

The capacity of entanglement depends on the choice of the original state of the system. In this section we study this dependence by considering various simple examples.

A. Simple n -qubit examples

We begin by studying simple n -qubit systems. Two-qubit systems were discussed in [52]. As a simplest example, consider the states

$$|\theta, \phi\rangle = \cos(\theta/2)|10\rangle + e^{i\phi} \sin(\theta/2)|01\rangle \quad (4.1)$$

and form a reduced density matrix by tracing over the other spin,

$$\rho_{\text{red}} = \text{diag}(\sin^2(\theta/2), \cos^2(\theta/2)). \quad (4.2)$$

The entanglement entropy grows monotonically towards the maximum, but the capacity of entanglement goes to zero both in the maximally entangled and separable state limits while reaching a maximum at a partially entangled state with $\cos^2(\theta/2) \approx 0.2885$ (Fig. 1).¹¹

Moving to a three-qubit system, we consider the subspace

$$|\theta, \phi\rangle = \cos \theta |001\rangle + \sin \theta (\cos \phi |010\rangle + \sin \phi |100\rangle) \quad (4.6)$$

and form a reduced density matrix by tracing over one spin. In this case, we again see that the capacity of entanglement goes to zero in the limit of maximal entanglement and

¹¹Reference [52] considered a slightly more general example, starting with the states

$$|\psi\rangle = a_1|00\rangle + a_2|01\rangle + a_3|10\rangle + a_4|11\rangle \quad (4.3)$$

with the normalization $\sum_{i=1}^4 |a_i|^2 = 1$. The diagonalized reduced density matrix for one spin takes the same form as above, but [52] wrote it in terms of the concurrence c of the 2-bit system,

$$\rho_{\text{red}} = \text{diag}\left(\frac{1 + \sqrt{1 - c^2}}{2}, \frac{1 - \sqrt{1 - c^2}}{2}\right), \quad (4.4)$$

where

$$c = 2|a_1 a_4 - a_2 a_3|. \quad (4.5)$$

They also observed that in general $\Delta S_{EE}^2 \neq S_{EE}$, equality holds only at $c = 0$ and at special value $c \approx 0.8272$.

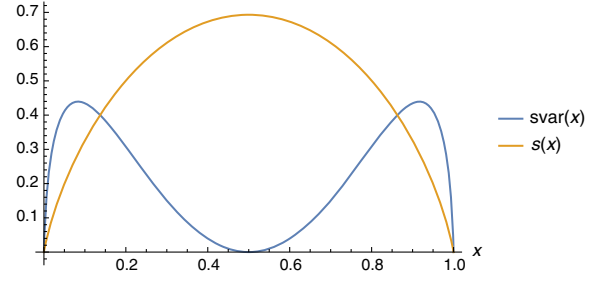


FIG. 1. Entanglement entropy and the capacity of entanglement for the state (4.2) as a function of $x = \cos^2(\theta/2)$.

separable states, while reaching a maximum for partially entangled states (Fig. 2).

As these simple examples illustrate, the entanglement entropy and its fluctuations, characterized by the cumulants have no reason to be equal, except for some special states. The capacity of entanglement characterizes the width of the eigenvalue spectrum of the density matrix. We study that by a simple example. Let us assume that the (reduced) system has N levels, and we have moved to a diagonal basis. We then assume that the N eigenvalues λ_i of the (reduced) density matrix have a gaussian distribution with variance σ , i.e., they form a Gaussian vector (an example is shown in Fig. 3).

We can now interpolate from $\sigma = 0$, where the density matrix describes a pure state, to $\sigma \rightarrow \infty$ limit giving a maximally mixed state. We then plot the (entanglement) von Neumann entropy and the capacity of entanglement $C_E = \Delta S^2$ as a function of σ in Fig. 4. The left panel (A) shows only the capacity of entanglement to highlight its non-trivial dependence of the spread of eigenvalues; the right panel (B) shows it together with the entropy.

B. Random bipartite entanglement

One could imagine that the origin of entanglement and capacity of entanglement has to do with the fact that we are projecting a random pure state to a subsystem. It is therefore interesting to compute the expectation value of the entanglement entropy and capacity of entanglement in such a setting. More precisely, we are going to consider random pure states in a Hilbert space of dimension pq of the form $\mathcal{H}_q \otimes \mathcal{H}_p$, projected to the subsystem with Hilbert space \mathcal{H}_p of dimension p ($q \geq p$), a review is e.g., [53]. In this way we will get a probability distribution of reduced density matrices, and we can compute the expectation value of entanglement entropy and the capacity of entanglement in this ensemble.

The probability distribution for the eigenvalues λ_i of the reduced density matrix is [54]

$$P_{p,q}(\lambda_1, \dots, \lambda_p) = \mathcal{N}_{p,q} \prod_{1 \leq i < j \leq p} (\lambda_i - \lambda_j)^2 \prod_{1 \leq k \leq p} \lambda_k^{q-p} \delta \times \left(1 - \sum_{i=1}^p \lambda_i\right). \quad (4.7)$$

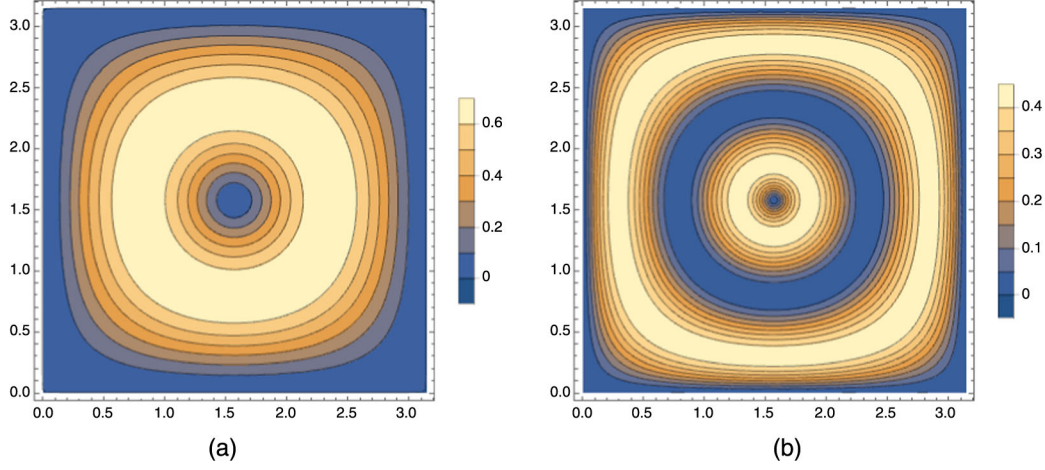


FIG. 2. Entanglement entropy (a) and the capacity of entanglement (b) plotted as a function of the angle parameters θ, ϕ of the state (4.6).

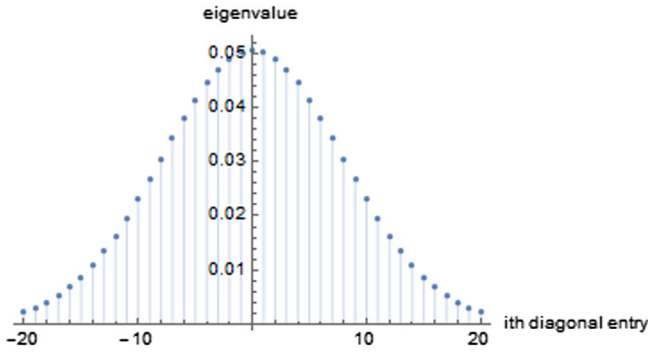


FIG. 3. $N = 41$ eigenvalues λ_i of ρ with a Gaussian profile. The indexing is $i = -20, -19, \dots, 20$.

where $\mathcal{N}_{p,q}$ is a known normalization factor [55]. To compute the expectation value of the Renyi entropies, we need to compute the expectation value of $\sum \lambda_i^\alpha$ under this probability distribution. The relevant integrals can be

extracted from the results in [56]. For example, we find that when $p = q = 2$

$$\langle \text{tr} \rho^\alpha \rangle_{2,2} = \mathcal{N} \frac{2(\alpha^2 + \alpha + 2)\Gamma(\alpha + 1)}{\Gamma(\alpha + 4)} \quad (4.8)$$

from which we obtain

$$\langle S_{EE} \rangle_{2,2} = \frac{1}{3}, \quad \langle C_E \rangle_{2,2} = \frac{13}{36}. \quad (4.9)$$

We therefore see that for $p = q = 2$ the entanglement entropy and capacity are very close to each other.

For arbitrary p, q an exact answer for $\langle S_{EE} \rangle_{p,q}$ was conjectured in [57] and proven in [58],

$$\langle S_{EE} \rangle_{p,q} = \sum_{k=q+1}^{pq} \frac{1}{k} - \frac{p-1}{2q}, \quad (4.10)$$

for $p \leq q$.

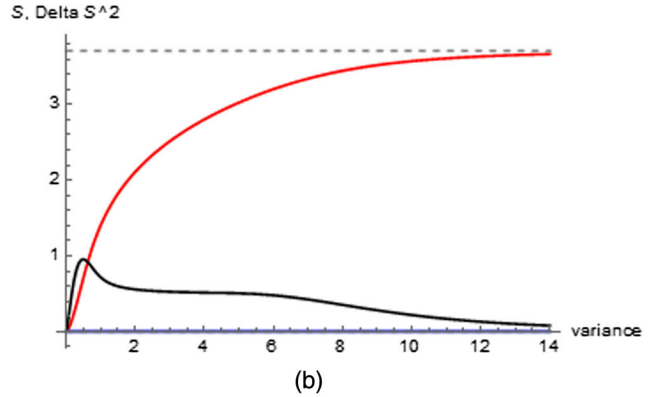
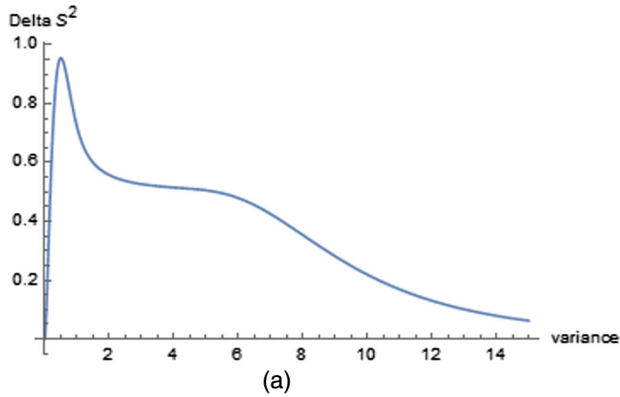


FIG. 4. (Entanglement) entropy and the capacity of entanglement plotted as a function of the variance σ of the eigenvalue distribution. The left panel (a) shows the capacity of entanglement $C_E = \Delta S^2$, the right panel (b) superimposes it (lower curve) with the entropy (upper curve). The dashed line indicates the maximum entropy.

Similarly we can compute

$$\langle \text{tr} \rho^\alpha \rangle_{2,q} = \mathcal{N} \frac{2(q-2)!(\alpha^2 + \alpha + 2q - 2)\Gamma(\alpha + q - 1)}{\Gamma(\alpha + 2q)} \quad (4.11)$$

from which one can obtain for example $\langle S_{EE} \rangle_{2,3} = \frac{9}{20}$ and $\langle C_E \rangle_{2,3} = \frac{1169}{3600}$. Numbers quickly become unwieldy, but as we increase p and q in general the entanglement entropy increases and the capacity divided by entanglement decreases. This is consistent with the idea that as p and q increase, the reduced density matrix becomes more and more maximally mixed. To illustrate, for $p = 2$ and $q = 100$ we get $\langle S_{EE} \rangle_{2,100} \simeq 0.686$, close to the maximal value $\log 2 \simeq 0.693$, while $\langle C_E \rangle_{2,100} \simeq 0.015$.

The case $p = q = 2$ is the case where entanglement entropy and capacity are nearly equal to each other. The ratio C_E/S_{EE} decreases even when we keep $p = q$ and increase p and q . Already for $p = q = 3$ where $S_{EE} = 1669/2520$ and $C_E = 2898541/6350400$ the ratio has decreased to $C_E/S_{EE} \simeq 0.689$. As we briefly review below, this can be proved rigorously by studying the limit of large p, q using random matrix theory [59,60].

The main observation of these studies is that a possible explanation for the approximate equality of S_{EE} and C_E in a system is that most of the entanglement is being carried by randomly entangled pairs of qubits (the case $p = q = 2$).

1. Random pure states and Wishart-Laguerre random matrices

We may also study the reduced density matrices of random pure states by using their connection to the Wishart-Laguerre ensemble of random matrices. We are not going to perform a mathematically rigorous analysis here (along the lines of [59]), but just quickly study whether the ratio C_E/S_{EE} decreases to zero in the limits $p = q \rightarrow \infty$, and $p, q \rightarrow \infty$ with $p/q \rightarrow 0$, following the approach in [60]. In [60], the reduced density matrix was written in the form

$$\rho = \frac{YY^\dagger}{\text{Tr}(YY^\dagger)}, \quad (4.12)$$

where $W \equiv YY^\dagger$ is a $p \times p$ random matrix belonging to the $\beta = 2$ Wishart-Laguerre ensemble. In the limit of large p, q with $\alpha = p/q$ fixed, one may replace $\text{Tr}(YY^\dagger) \rightarrow p$, so that the eigenvalues λ_i of ρ are related to the eigenvalues μ_i of the Wishart matrix W by a simple rescaling

$$\lambda_i = \frac{\mu_i}{p}. \quad (4.13)$$

This gives a quick way of computing the ensemble average of the Rényi entropy, by using the average spectral density of the Wishart ensemble. We start with

$$k(s) = \text{Tr}(\rho^s) = \sum_i \lambda_i^s \approx \frac{1}{p^s} \int d\mu \sum_i \delta(\mu - \mu_i) \mu^s \quad (4.14)$$

where

$$\sum_i \delta(\mu - \mu_i) \equiv \hat{\nu}(\mu) \quad (4.15)$$

is the spectral density of the Wishart matrix. Its ensemble average $\langle \hat{\nu}(\mu) \rangle$ is given by the Marchenko-Pastur distribution [61] $\bar{n}(\mu)$,

$$\langle \hat{\nu}(\mu) \rangle = p \bar{n}(\mu) = \frac{p}{2\pi\alpha\mu} \sqrt{(\mu - \alpha_-)(\alpha_+ - \mu)}, \quad (4.16)$$

where $\alpha_- \leq \mu \leq \alpha_+$, with $\alpha_\pm = (1 \pm \sqrt{\alpha})^2$, $\alpha = p/q$. The ensemble average $\langle k(s) \rangle$ thus becomes

$$\begin{aligned} \langle k(s) \rangle &= \langle \text{Tr} \rho^s \rangle = \frac{1}{p^{s-1}} \int d\mu \bar{n}(\mu) \mu^s \\ &= p^{1-s} \frac{1}{2\pi\alpha} \int_0^{\alpha_+ - \alpha_-} dy \sqrt{y(\alpha_+ - \alpha_- - y)} (y + \alpha_-)^{s-1} \\ &= \left(\frac{p}{\alpha_-} \right)^{1-s} {}_2F_1 \left(1-s, \frac{3}{2}, 3, \frac{-4\sqrt{\alpha}}{\alpha_-} \right). \end{aligned} \quad (4.17)$$

The general result from (4.17) for the entanglement entropy and the capacity of entanglement is then

$$\begin{aligned} \langle S_{EE} \rangle &= \log \frac{p}{\alpha_-} + G_a^{(1)} \left(0, \frac{3}{2}, 3, \frac{-4\sqrt{\alpha}}{\alpha_-} \right) \\ \langle C_E \rangle &= G_a^{(2)} \left(0, \frac{3}{2}, 3, \frac{-4\sqrt{\alpha}}{\alpha_-} \right) - \left(G_a^{(1)} \left(0, \frac{3}{2}, 3, \frac{-4\sqrt{\alpha}}{\alpha_-} \right) \right)^2 \end{aligned} \quad (4.18)$$

where¹²

$$G_a^{(n)}(a, b, c, z) = \frac{\partial^n}{\partial a^n} {}_2F_1(a, b, c, z). \quad (4.19)$$

In the limit $\alpha = p/q \rightarrow 0$, one obtains

$$\frac{\langle C_E \rangle}{\langle S_{EE} \rangle} \approx \frac{\frac{p}{q}}{\log p - \frac{p}{2q}} \rightarrow 0, \quad (4.20)$$

confirming the numerical guess. In the limit $p = q$ (4.17) gives

¹²Properties of these derivatives of the hypergeometric function can be found in [62]. They are also built into the *Mathematica* program.

$$\begin{aligned}\langle k(s) \rangle &= p^{1-s} \frac{1}{2\pi} \int_0^4 dy \sqrt{y(4-y)} y^{s-1} \\ &= \frac{4}{\sqrt{\pi}} \left(\frac{p}{4}\right)^{1-s} \frac{\Gamma(s+\frac{1}{2})}{\Gamma(s+2)},\end{aligned}\quad (4.21)$$

in agreement with Eq. (12) in [59], from which we obtain

$$\langle S_{EE} \rangle \approx \log p - \frac{1}{2}; \quad \langle C_E \rangle \approx \frac{\pi^2}{3} - \frac{11}{4} \approx 0.5399 \quad (4.22)$$

also confirming the numerical guess $\langle C_E \rangle / \langle S_{EE} \rangle \rightarrow 0$ in the limit $p = q \rightarrow \infty$.

We therefore see that a possible explanation for approximate equality of S_{EE} and C_E in a system is that most of the entanglement is being carried by randomly entangled pairs of qubits.

V. THE CAPACITY OF ENTANGLEMENT IN QUANTUM FIELD THEORIES AND UNIVERSALITY

A. 1 + 1 dimensional CFTs

We now move to consider conformal field theories in various dimensions. In particular, in 1 + 1 dimensions we can make use of very general analytical results for the Rényi entropies. Rényi entropies for generic CFTs have been computed in a variety of cases: for an initial vacuum state for infinite and finite systems, for thermal backgrounds for finite and infinite systems, and for non-equilibrium dynamics involving different quench protocols, notably global and local quenches.¹³ A recent review is e.g., [64]. The computations lead to

$$\text{Tr}(\rho_A^\alpha) = c_\alpha e^{-\frac{c}{12}(\alpha-\frac{1}{\alpha})W_A}, \quad (5.1)$$

where W_A is a function of the size of the subsystem (and time, for quenched systems) but independent of α , and c_α with $c_1 = 1$ are model specific constants that do not depend on the size l . The Rényi entropies $S_\alpha(A)$ are then

$$S_\alpha(A) = \frac{c}{12} \left(1 + \frac{1}{\alpha}\right) W_A + (1 - \alpha)^{-1} \log c_\alpha. \quad (5.2)$$

From the cumulant generating function $\tilde{k}(\alpha) = (1 - \alpha)S_\alpha$, we immediately obtain the general result for the capacity of entanglement,

$$C_E = S_{EE} + c'_1 - (c'_1)^2 + c''_1 = \frac{c}{6} W_A + c'_1 + (c'_1)^2 + c''_1, \quad (5.3)$$

¹³In Sec. II we discussed the relation of the Rényi entropies and the entanglement spectrum. A recent paper [63] studies the time evolution of the entanglement spectrum and the entanglement Hamiltonian after a quench.

and in particular, for large W_A , the last three terms on the right-hand side are negligible. Explicit results for W_A for various different situations are listed in [65]. For time independent cases with A a finite interval of length l ,

$$W_A = \begin{cases} 2 \log(\frac{l}{\epsilon}) & \text{infinite system, } T = 0 \\ 2 \log(\frac{L}{\pi\epsilon} \sin \frac{\pi l}{L}) & \text{finite system, size } L, T = 0 \\ 2 \log(\frac{\beta}{\pi\epsilon} \sinh \frac{\pi l}{\beta}) & \text{infinite system, } T > 0. \end{cases} \quad (5.4)$$

For time-dependent cases with two semi-infinite intervals, some known results are

$$W_A = \begin{cases} \log(\frac{\beta}{\pi\epsilon} \cosh(2\pi t/\beta)) & \text{global quench at } t = 0 \\ \log(\frac{l^2 + \lambda^2}{4\epsilon^2}) & \text{local quench at } t = 0, \end{cases} \quad (5.5)$$

with quench parameters β (the resulting inverse temperature) and λ .

It may be somewhat surprising that the leading order result $C_E = S_{EE}$ holds in all of the above cases, even for the quenched nonequilibrium systems throughout their time evolution. The entangled states are rather special, prepared by using conformal mappings. Also, the result tells us that after the quench, the growth rate of the capacity of entanglement is the same as that of the entanglement entropy. The growth of the entanglement entropy is consistent with the interpretation where the spreading of entanglement is carried by quasiparticle pairs created at the quench and propagating ballistically through the system at the speed of light. It is not obvious at all that this picture should imply the same growth rate for the capacity of entanglement. Consider for example the following toy model. Suppose we quench a system and afterwards entanglement builds up because pairs of perfectly entangled quasiparticles are created. The reduced density matrix of a subsystem will then roughly consist of the original reduced density matrix combined with say $n(t)$ qubits which are pairwise perfectly entangled with $n(t)$ qubits outside the subsystem. The number of such qubits will be time dependent. We will model the reduced density matrix as

$$\rho(t) = \rho_{UV} \otimes \left(\frac{1}{2^{n(t)}} \mathbb{I}_{2^{n(t)} \times 2^{n(t)}} \right) \otimes |\psi(t)\rangle \langle \psi(t)| \quad (5.6)$$

where ρ_{UV} represents the very UV d.o.f. which were entangled, and remain entangled across the boundary of the region. The pure state $\psi(t)$ represents the reservoir of states from which qubits are extracted as more and more entangled quasiparticles become relevant. This pure state lives in some Hilbert space of dimension $D - 2^{n(t)}$ with a very large D . Details of the reservoir pure state are irrelevant.

It is straightforward to compute the Rényi entropies for this density matrix,

$$\frac{1}{1-\alpha} \log \text{Tr}(\rho(t)^\alpha) = S_{\text{UV}}^\alpha + n(t) \log 2. \quad (5.7)$$

We then deduce that EE grows as a function of t ,

$$S_{EE} = S_{EE}^{\text{UV}} + n(t) \ln 2, \quad (5.8)$$

where the function $n(t)$ can be adjusted to the quasiballistic interpretation (with the characteristic long linear growth regime), but the capacity of entanglement stays constant

$$C_E = \Delta S_{\text{UV}}^2 \quad (5.9)$$

receiving contributions only from the UV, not from the created quasiparticles. Therefore if the capacity of entanglement and EE grow identically after the quench, it is in tension with the simple ballistic propagation of created pairwise perfectly entangled quasiparticles which leads to the above model.

Rather than expecting the quasiparticles to be pairwise perfectly entangled, we could expect them to be created in random pure states with some statistical distribution. In Sec. IV B we studied random pure states with bipartite entanglement between d.o.f. inside and outside the subsystem. There we saw that it was indeed possible for the capacity of entanglement to be approximately equal to the entanglement entropy, for ensemble averages of randomly entangled pairs of qubits. On the other hand, in Sec. III B we saw that such an equality also arises (in the large N limit) from a more detailed connection between particle fluctuations into a subregion and its entropy. We conclude that the intuitive interpretation of entanglement being carried by ballistic propagation quasiparticles needs to be refined with correct statistics for the distribution of entanglement.

Finally, we make some brief comments on the two interval case. Let the two intervals be $A = [u_1, v_1]$ and $B = [u_2, v_2]$. The Rényi entropy for the two interval case has the general form [66]

$$\text{Tr} \rho_{A \cup B}^\alpha = c_\alpha \left(\frac{(u_2 - u_1)(v_2 - v_1)}{|(v_2 - u_1)(v_2 - u_2)(v_1 - u_1)(v_1 - u_2)|} \right)^{(c/6)(\alpha - \frac{1}{\alpha})} \times \mathcal{F}_\alpha(\eta) \quad (5.10)$$

where c_α is a nonuniversal constant and \mathcal{F}_α is a nonuniversal function of the cross ratio η . This nonuniversal function makes the analysis more complicated, so here we consider only the case $\mathcal{F}_\alpha = 1$, such as in the massless fermion theory [67–69]. Then the Rényi entropy takes the same form as (5.1), and we readily obtain a similar relationship for the leading terms as in (5.3),

$$C_E(A \cup B) = S_{EE}(A \cup B). \quad (5.11)$$

B. An alternative calculation in higher dimensions

In subsection III E we discussed an alternative calculation to compute the capacity of entanglement for a $1+1$ dimensional conformal field theory. In this section we discuss its generalization to higher dimensional CFTs.

Suppose we consider the entanglement of a ball. Inserting the explicit form of the modular Hamiltonian, we obtain a suitably integrated two-point function of the energy-momentum tensor, in agreement with (3.40) and (3.41). Such types of integrated correlation functions were studied in [24,25]. The idea of the computation is as follows. A ball in a constant timeslice is mapped to itself under a subgroup $SO(1, d-1)$ of the conformal group, and the double integral is invariant under the action of this conformal group. We can use this symmetry to fix one of the points at the origin, leaving us with a single integral over the ball. This integral is still divergent, but because the modular Hamiltonian involves the integral of a conserved current over a fixed timeslice, we can evaluate it on any suitable time slice, and in particular we can move the slice a little bit so as to avoid the origin. What is then left is a finite integral, which we still need to multiply by the volume of the gauge group that we fixed. This volume is the volume of $SO(1, d-1)/SO(d-1)$, which is the Euclidean hyperbolic space, because the origin is left fixed by a $SO(d-1)$ subgroup of $SO(1, d-1)$. This hyperbolic space is the same space that appears at the boundary of AdS if we map the Rindler wedge associated to the ball to a hyperbolic black hole as in [40].

It is clear from the above that the final result will only depend on the coefficient C_T which appears in the two-point function of two energy-momentum tensors

$$\langle T(x)T(0) \rangle \sim \frac{C_T}{x^{2d}}. \quad (5.12)$$

The relevant computation of the integrated two-point function was studied in detail in [24] for arbitrary CFT's, and Eq. (1.3) therein reads

$$S'_{q=1} = -\text{Vol} \mathbb{H}^{d-1} \cdot \frac{\pi^{d/2+1} \Gamma(d/2)(d-1)}{(d+1)!} C_T \quad (5.13)$$

where the volume of hyperbolic space appears as we just explained, and we can interpret the result as the capacity of entanglement by $C_E = \Delta S_{EE}^2 = -2S'_{q=1}$. The result is indeed proportional to the coefficient C_T that appears in the two-point function of two energy-momentum tensors. For later use, we write the result in a different form. As also explained in [24], the C_T above should not be confused with the coefficient c appearing in the $d=4$ trace anomaly of a CFT or the \tilde{C}_T used in [41] in their expression for the Rényi entropies. They are proportional to each other,

$$\tilde{C}_T = \frac{\pi^d (d-1)}{(d+1)!} C_T \quad (5.14)$$

such that $\tilde{C}_T|_{d=4} = c$. Thus, from (5.13), we arrive the expression

$$C_E = \tilde{C}_T \frac{2\Gamma(d/2)}{\pi^{d/2-1}} \text{Vol}\mathbb{H}^{d-1}. \quad (5.15)$$

This expression is similar to that of the entanglement entropy. When computing the entanglement entropy via the conformal mapping to hyperbolic space, the original computation in [40] obtained the expression

$$S_{EE} = a_d^* \frac{2\Gamma(d/2)}{\pi^{d/2-1}} \text{Vol}\mathbb{H}^{d-1}, \quad (5.16)$$

where a_d^* is a dimensionless constant, a central charge. In even dimensional spacetimes, this equals the coefficient of the A-type trace anomaly of the CFT. In odd dimensional spacetimes, $a_d^* \propto \log Z_{S^d}$, i.e., the logarithm of the partition function of the CFT on a unit sphere.

The regulated hyperbolic volume is given explicitly in (3.1) in [24] and is also easy to compute independently. The expression is sensitive to the choice of the regularization scheme, but even and odd d have a universal coefficient for the logarithmic and constant term, respectively. They are

$$\text{Vol}\mathbb{H}^{d-1} = \frac{\pi^{d/2}}{\Gamma(d/2)} \begin{cases} (-)^{d/2-1} 2\pi^{-1} \log(R/\varepsilon) & \text{d even} \\ (-)^{(d-1)/2} & \text{d odd} \end{cases}. \quad (5.17)$$

To conclude, the above computations suggest a simple universal ratio for the entanglement entropy and the capacity of entanglement,

$$\frac{C_E}{S_{EE}} = \frac{\tilde{C}_T}{a_d^*}. \quad (5.18)$$

We discuss universality in more detail in Sec. V F.

C. The entanglement spectrum

The capacity of entanglement captures a particular feature of the full entanglement spectrum. Here we briefly review some aspects of the full entanglement spectrum and comment on the implications for the capacity of entanglement.

We start with the two-dimensional case where the entanglement spectrum can be explicitly obtained in several cases. In particular, from (5.1) one can in principle extract the entanglement spectrum for $1+1$ dimensional CFT's using an inverse Laplace transform. Of course, the precise answer will depend on the detailed structure of c_α , but it is instructive to show the spectrum in an explicit example. For $c_\alpha = 1/\alpha$, the inverse Laplace transform leads to the following result for the density of eigenvalues of the reduced density matrix [10,16]

$$\rho(\lambda) = \theta(-\log \lambda - cW_A/12)\lambda^{-1} \times I_0\left(2\sqrt{cW_A(-\log \lambda - W_A)/12}\right) \quad (5.19)$$

and one can verify explicitly that indeed

$$\int_0^1 \rho(\lambda)\lambda^\alpha = \alpha^{-1} e^{-\frac{c}{12}(\alpha-\frac{1}{\alpha})W_A}. \quad (5.20)$$

It is perhaps more instructive to rewrite this in terms of an energy spectrum with $\lambda = e^{-E}$. If we denote

$$E_c = \frac{c}{12} W_A \quad (5.21)$$

then

$$\rho(E) = \theta(E - E_c) I_0\left(2\sqrt{E_c(E - E_c)}\right) \quad (5.22)$$

and indeed

$$\int_0^\infty dE \rho(E) e^{-\alpha E} = \alpha^{-1} e^{-\frac{c}{12}(\alpha-\frac{1}{\alpha})W_A}. \quad (5.23)$$

It is interesting to see that the spectrum cuts off at a value E_c which is determined by the UV cutoff that one needs to introduce in order to regulate the divergences in the computation of the Rényi entropies. One is almost tempted to view E_c as some sort of analog of the Casimir energy. As we remove the cutoff, all eigenvalues of the density matrix are located in an ever smaller interval starting at $\lambda = 0$. In the above, we only considered the leading term of the Rényi entropy in terms of a cutoff, one can in principle do more precise computations taking the exact cutoff dependence into account (which involves a choice of boundary state) [13,16], but that is beyond the scope of this paper.

Ultimately, in $1+1$ dimensions, (5.22) follows from conformal invariance, but in a rather indirect way. It would be desirable to have a more direct understanding of the connection between conformal invariance and the various features of (5.22), as that might shed further light on the relation between C_E and S_{EE} in higher dimensions.

Expressions where c_α is some other integer power of α can in principle be obtained from (5.22) by differentiating or integrating with respect to E . The feature of (5.22) which does not depend much on these details is its large E behavior. For large z , the asymptotic behavior of $I_\nu(z)$ is $I_\nu(z) \sim e^z/\sqrt{2\pi z}$. Keeping only the exponential we therefore see that $\rho(E)$ behaves as

$$\rho(E) \sim \exp\left(2\sqrt{E_c(E - E_c)}\right), \quad E \rightarrow \infty. \quad (5.24)$$

Another interpretation of this expression is that it is the density of states for a $1+1$ dimensional CFT on the plane. In higher dimensions, it maps to the density of states on the

hyperbolic plane, but the one-dimensional hyperbolic plane is just the ordinary line. Actually, according to [40] after a change of coordinates the causal diamond attached to the spatial interval $r \in [-l/2, l/2]$ maps under the change of coordinates

$$t \pm r = \tanh \left(\frac{1}{2} \left(\frac{\tau}{R} \pm u \right) \right) \quad (5.25)$$

to

$$ds^2 = \frac{-d\tau^2 + R^2 du^2}{(\cosh u + \cosh(\tau/R))^2} \quad (5.26)$$

where, for an interval of length l , $R = l/2$. Putting a UV cutoff at $r = \pm(R - \epsilon)$ corresponds under this change of coordinates to $u \simeq \pm \log(l/\epsilon)$ for small ϵ . The length of the u -interval is therefore precisely what we called W_A .

We therefore need to compute the density of states on the plane for a box of size W_A and relate the energy on the plane to the energy on the causal diamond. The change of coordinates (5.25) has a Schwarzian derivative of $-1/2$. To get the full shift in the energy we need to integrate this over the spatial box. Therefore the full shift in the left- or right-moving energy is $cW_A/24$. This leads to a shift in the total energy of the form

$$E_{\text{plane}} \sim E_{\text{diamond}} - \frac{c}{12} W_A. \quad (5.27)$$

Moreover, the energy spectrum for a box of size W_A is quantized in units of $\sim 2\pi/W_A$, and there is an extra factor of 2π relating the energy to L_0 , so altogether what one should use in the Cardy formula is

$$L_0^{\text{plane}} \sim \frac{W_A}{(2\pi)^2} \left(E_{\text{diamond}} - \frac{c}{12} W_A \right) \quad (5.28)$$

which explains the form of (5.24) from the usual Cardy formula.

In higher dimensions, we can study features of the entanglement spectrum whenever the CFT has a holographic dual, since the Rényi entropies can then be obtained from the thermodynamics of hyperbolic black holes [41]. The density of states will now obey

$$\int_0^\infty dE \rho(E) e^{-\alpha E} = e^{-f(\alpha)} \quad (5.29)$$

where

$$f(\alpha) = \alpha S_{EE} \left(1 - \frac{x_\alpha^d}{2} - \frac{x_\alpha^{d-2}}{2} \right) \quad (5.30)$$

with

$$x_\alpha = \frac{1 + \sqrt{1 + \alpha^2 d(d-2)}}{\alpha d}. \quad (5.31)$$

We notice that as $\alpha \rightarrow \infty$, $f(\alpha) \sim E_c \alpha$, where

$$E_c = S_{EE} \left(1 - \left(\frac{d-2}{d} \right)^{\frac{d}{2}} - \left(\frac{d-2}{d} \right)^{\frac{d-1}{2}} \right) \quad (5.32)$$

which implies in particular that $\rho(E) = 0$ for $E < E_c$. It is tempting to interpret E_c once more as some sort of Casimir energy. The high-energy behavior of ρ is determined by the behavior of $f(\alpha)$ as $\alpha \rightarrow 0$ which is $f(\alpha) \sim c_d S_{EE} \alpha^{1-d}$ from which it follows that for large E

$$\log \rho(E) \sim c'_d E_c^{1/d} (E - E_c)^{(d-1)/d} \quad (5.33)$$

where c'_d is some constant which only depends on the dimension d . The scaling with E is exactly what one expects generically for a CFT at high energy/temperature, this simply follows from extensivity of the free energy. Corrections to (5.33) can be obtained from subleading terms in the expansion near $\alpha = 0$. The first subleading term scales as α^{3-d} which gives rise to corrections to (5.33) of the form

$$\log \rho(E)_{\text{subleading}} \sim c''_d E_c^{3/d} (E - E_c)^{(d-3)/d} \quad (5.34)$$

It is not clear whether these results also hold for more general CFT's without a holographic dual. We also observe that the capacity of entanglement is related to the behavior of $f(\alpha)$ near $\alpha = 1$, whereas the cutoff energy E_c and the high-energy behavior are related to the large and small α behavior of $f(\alpha)$. *A priori*, these are unrelated to each other, and therefore the capacity of entanglement does not in general make any prediction for the behavior of the density of states at either high or low energy. If the theory has a holographic dual we know the explicit form of $f(\alpha)$, which only depends on the dimension d , and there are some simple factors of order unity which relate the behavior for small α , α of order 1, and large α . We also confirm once more that $C_E = S_{EE}$ in because $C_E = -f''(1) = S_{EE} = f'(1)$ by explicit computation.

We notice that in the above cases not just the capacity of entanglement, but all Rényi entropies obey an area law. This implies that the function $f(\alpha)$ in (5.29) has an area law and scales as a^{2-d} where a is a short-distance UV cutoff. This translates into the following leading cutoff dependence of the density of states

$$\log \rho(E) \sim \frac{1}{a^{d-2}} \log \hat{\rho}(E a^{d-2}) + \dots \quad (5.35)$$

where $\hat{\rho}$ does not depend on the cutoff. This is precisely the scaling one would get from a local Rindler point of view. Therefore, the area law for capacity of entanglement (and more generally for the Rényi entropies) seems to arise from the fact that these quantities are dominated by UV d.o.f. localized near the entangling surface, which in turn can be well approximated by local Rindler modes.

D. The capacity of entanglement in holographic systems

Entanglement Rényi entropy was considered for systems with a holographic dual with a black hole in [41]. Their method was to relate Rényi entropies of spheres to thermal entropies in hyperbolic spacetime at a range of temperatures. This method is applicable only for systems with bulk duals that admit black hole solutions at various temperatures. After obtaining the Rényi entropies, the computation of entanglement entropy and the capacity of entanglement are straightforward. All the holographic Rényi entropies are proportional to the volume of hyperbolic space, which contains all the divergent terms. As discussed in [41], the ratio S_a/S_1 of the Rényi entropies then yields universal information characterizing the dual CFTs. In similar spirit, we choose to study the ratio of the entanglement entropy and the capacity of entanglement, another universal constant (at least for spherical entangling surfaces).

In holographic duals of Einstein gravity, the entanglement entropy and the capacity of entanglement are equal,

$$S_{EE} = 2\pi \left(\frac{\tilde{L}}{l_p} \right)^{d-1} \text{Vol}\mathbb{H}^{d-1} = C_E, \quad (5.36)$$

where $\text{Vol}\mathbb{H}^{d-1}$ denotes the volume of the hyperbolic space. Here \tilde{L} is the curvature scale of the dual AdS spacetime. In Einstein gravity, it equals the curvature scale in the cosmological constant term but it is not the case for higher curvature theories.

In Gauss-Bonnet gravity, the ratio becomes more interesting. First, for $d = 4$, the bulk and boundary theories have two central charges, a and c , that appear in the boundary Weyl anomaly. The entanglement entropy and the capacity of entanglement are

$$S_{EE} = \frac{2a}{\pi} \text{Vol}\mathbb{H}^{d-1}, \quad C_E = \frac{2c}{\pi} \text{Vol}\mathbb{H}^{d-1}, \quad (5.37)$$

so that in $d = 4$ the gravity calculation indeed produces the same ratio (5.18) as the field theory calculation. *A priori*, one might not have expected any constraints for the ratio of the two. However, the ratio of the CFT central charges is restricted by the “conformal collider bounds” proposed in [26] and proven in [27], implying a bound

$$\frac{18}{31} \leq \frac{C_E}{S_{EE}} = \frac{c}{a} \leq 3, \quad (5.38)$$

for unitary 4d CFTs with gravity duals. The central charges a and c are known explicitly for some field theories. It is interesting to compute the ratio even when the theories are not holographic—one may ask if the holographic “prediction” (5.38) still holds. For example, for conformal free scalar field theory and massless free Dirac fermion theory the ratios of c and a are [70]

$$\frac{c_{\text{scalar}}}{a_{\text{scalar}}} = 3, \quad \frac{c_{\text{fermi}}}{a_{\text{fermi}}} = \frac{18}{11}. \quad (5.39)$$

For general $d \geq 4$ the theory is parametrized with λ that is the coefficient for the 4-dimensional Euler density term [41],

$$I = \frac{1}{2l_p^{d-1}} \int d^{d+1}x \sqrt{-g} \left[\frac{d(d-1)}{L^2} + R + \frac{\lambda L^2}{(d-2)(d-3)} \chi^4 \right]. \quad (5.40)$$

In this case one finds

$$S_{EE} = 2\pi \left(1 + \frac{(d-1)(-1 + \sqrt{1-4\lambda})}{d-3} \right) \left(\frac{\tilde{L}}{l_p} \right)^{d-1} \text{Vol}\mathbb{H}^{d-1}, \quad (5.41)$$

$$C_E = 2\pi \sqrt{1-4\lambda} \left(\frac{\tilde{L}}{l_p} \right)^{d-1} \text{Vol}\mathbb{H}^{d-1} \quad (5.42)$$

so that the ratio C_E/S_{EE} is again a universal λ - and d -dependent constant. Here, \tilde{L} is proportional to the L in the action, $\tilde{L} = L\sqrt{2\lambda}/(\sqrt{1-\sqrt{1-4\lambda}})$.

E. Violation of the area law

It has been known for some time that the area law of entanglement entropy is violated in some systems, we point out that the same is true for the capacity of entanglement. Apart from the logarithmic scaling in $1+1$ dimensional conformal field theories, the most common example is provided by critical systems with a finite Fermi surface such as free fermions or Fermi liquids [71–74]. The Rényi entanglement entropy of these systems has been computed to the leading order in [75]

$$S_\alpha = \left(1 + \frac{1}{\alpha} \right) \frac{1}{(2\pi)} \frac{1}{24} \int_k \int_x dA_k dA_x |n_x \cdot n_k| \log \left(\frac{L}{\epsilon} \right) \quad (5.43)$$

where the integrals are taken over the entangling surface and Fermi surface, the n ’s are their normal unit vectors and L corresponds to the effective system length. The capacity of entanglement again tracks the entanglement entropy, with $C_E = S_{EE}$, both scaling as $L^{d-1} \log(L)$.

F. On the universality of the ratio of entanglement entropy and the capacity of entanglement

We have already seen in many field theories that the leading order of divergence of both the entanglement entropy and the capacity of entanglement are the same. It is interesting to ask if the ratio of the leading terms, computed using the same regularization scheme, is universal or if it is scheme dependent. At first thought, there is no immediate reason why there would not be some universality. After all,

they have the same power dependence on the regularization parameter. However, the choice of the regularization scheme could contain some hidden dependence on the quantity we are computing, spoiling the universality. We demonstrate this below by two cases, free massless scalar fields and fermions.

First, we consider the free massless scalar field theory with both $d = 3$ and $d = 4$. The Hamiltonian of the theory is

$$H = \frac{1}{2} \int_{\mathbb{R}^d} d^d x [(\pi)^2(x) + (\nabla \phi)^2(x)], \quad (5.44)$$

with the standard commutation relations. The entangling surface is a sphere. As the first way of regularizing the theory, we expand the scalar field ϕ and its conjugate momentum field π in Fourier modes for $d = 3$ (and in spherical harmonics for $d = 4$), and discretize the remaining radial integral to a sum. The entanglement entropy and variance can be expressed in terms of correlation matrices inside the sphere. This method of computation was originally used in [76] and a more detailed explanation can be found in [77].

Using these methods, we performed a quick numerical computation of the entanglement entropy and the capacity of entanglement at various radii and found that

$$\frac{C_E}{S_{EE}} \approx \begin{cases} 3 & d = 3 \\ 5 & d = 4 \end{cases}. \quad (5.45)$$

These ratios are only approximate. The Rényi entropy of this model in the four dimensional case was considered to greater accuracy in [78] using the same regularization scheme and can be used to confirm our approximate ratio.

As the second alternative regularization scheme we consider the heat kernel method. In this way, the Rényi entropy of the free massless scalar field theory for a spherical region was also computed analytically in [79]. The effective action, $W_n = -\log Z_n$, of the n -sheeted spacetime is

$$W_n = -\frac{1}{2} \int_{\epsilon^2}^{\infty} \frac{ds}{s} \text{Tr} K(s, n), \quad (5.46)$$

where the regularization is done by limiting the lower limit of the integral. The heat kernel is defined as

$$K(s, n, X, X') = {}_n\langle X | e^{-s\mathcal{D}} | X' \rangle_n, \quad (5.47)$$

where \mathcal{D} is the field operator of the theory. The trace itself is

$$\text{Tr} K(s, n) = \frac{1}{(4\pi s)^{d/2}} \left(nV + \pi \frac{(1-n^2)}{3n} sA(\Sigma) \right). \quad (5.48)$$

With these, it can be seen that the ratio

$$\frac{C_E}{S_{EE}} = 1 \quad (5.49)$$

for the leading terms, exactly, for all d . This is an obvious disagreement with the previous result (5.45). We conclude that at least for the free massless scalar field theory, the ratio of the leading terms is scheme dependent.

While the universality fails for generic field theories, we may restrict our consideration to conformal field theories. An alternative regularization scheme valid for all Rényi entropies of conformal field theories is provided by the holographic computation technique used in [40,41] and discussed above. We can test universality by the following criterion. We separately expand the entanglement entropy S_{EE} and the heat capacity C_E , and then compare one by one the ratios of the leading terms and the ratios of subleading terms at the same order in the expansion. If universality holds, the ratios of respective terms in the expansions should all be the same in all regularization schemes. Conversely, if we compute in one regularization scheme the ratio of say, the leading terms, and in a different regularization scheme the ratio of subleading terms (at same order), and find a different ratio in the two schemes, universality does not hold. We show an example below.

Let us consider free massless fermions in $d = 3$, a conformal field theory. For the leading terms, we first perform a numerical computation using similar methods as above, specified in [77]. The numerical computation leads to

$$\frac{C_E}{S_{EE}} \approx 2.9, \quad (5.50)$$

which implies that the ratio of the leading terms is approximately 2.9.

Next we consider subleading terms, and a different regularization scheme. For $d = 3$, the expansion of Rényi entropies contains a universal constant term, related to topological properties of the system. For entanglement entropy, this is known as the F -term (or the negation of it). The F -term has been shown to be equal to the constant term in the free energy of the system restricted to a unit sphere, S^3 . This has been computed for free massless fermions in e.g., [80]. The corresponding term for the heat capacity has been computed e.g., in [24]. The ratio of these two terms is approximately 1.4, in disagreement with the ratio of leading terms.

Therefore, we can conclude that at least for generic conformal field theories, the ratio of coefficients in the expansions of entanglement entropy and capacity of entanglement are not in general universal.¹⁴

Narrowing the set of theories further, we could consider CFTs which have a gravity dual. In these cases, geometry may provide a natural regularization scheme. Indeed, in Secs. V B and V D we saw that both the entanglement entropy and the

¹⁴There is an exception: the universal coefficients in the expansions are scheme-independent, so there ratios should also be.

capacity of entanglement were proportional to the hyperbolic volume factor containing the divergences, leading to a clean ratio of the two. Regarding the interesting question which CFTs then have gravitational duals, we are lead to speculate that perhaps such natural ratios in CFTs are a hint of a dual gravitational interpretation.

G. On the shape dependence of the capacity of entanglement

Much of the discussion has focused on spherical or planar entangling surfaces. As a natural generalization, we briefly comment on the shape dependence of the capacity of entanglement.

For general shapes in $d > 2$ quantum field theories, the coefficient of the leading divergent term of the Rényi entropies is proportional to the area of the entangling surface and non-universal. The subleading terms have more variety and, in general, the expansion is different from the spherical case. Much work has been carried out to understand the shape dependence of entanglement Rényi entropies, and some universal results are known.

A well-studied case is the 4-dimensional CFTs, for which the universal log term can be expressed in terms of integrals over the entangling surface Σ , [81]

$$S_a^{\text{univ}} = - \left(\frac{f_a(\alpha)}{2\pi} \mathcal{R}_\Sigma + \frac{f_b(\alpha)}{2\pi} \mathcal{K}_\Sigma - \frac{f_c(\alpha)}{2\pi} \mathcal{C}_\Sigma \right) \log \left(\frac{1}{\epsilon} \right), \quad (5.51)$$

where

$$\begin{aligned} \mathcal{R}_\Sigma &= \int_\Sigma d^2y \sqrt{h} R_\Sigma, & \mathcal{C}_\Sigma &= \int_\Sigma d^2y \sqrt{h} C_{ab}^{ab}, \\ \mathcal{K}_\Sigma &= \int_\Sigma d^2y \sqrt{h} \left[\text{tr} K^2 - \frac{1}{2} (\text{tr} K)^2 \right]. \end{aligned} \quad (5.52)$$

Here y and h are the coordinates of the entangling surface and the induced metric, respectively, and R_Σ , K , and, C_{ab}^{ab} , the intrinsic Ricci scalar of Σ , the extrinsic curvature of Σ , and the contraction of the Weyl tensor along coordinates orthogonal to the entangling surface. The functions f depend only on the physical data of the CFT and n . The function f_a can be computed by mapping the sphere to hyperbolic spacetime. On the other hand, f_c and f_b can be studied using deformations of spherical entangling surfaces to first and second order, respectively [82]. In general, it is known that $f_a(1) = a$, $f_b(1) = f_c(1) = c$. The universal log term of the entanglement entropy is thus

$$S_{EE}^{\text{univ}} = - \frac{c}{2\pi} \left(\frac{a}{c} \mathcal{R}_\Sigma + \mathcal{K}_\Sigma - \mathcal{C}_\Sigma \right) \log \left(\frac{1}{\epsilon} \right). \quad (5.53)$$

In comparison, the universal log term of the capacity of entanglement for holographic theories is [82]

$$C_E^{\text{univ}} = 2 \left(- \frac{c}{4\pi} \mathcal{R}_\Sigma + \frac{f'_b(1)}{2\pi} \mathcal{K}_\Sigma - \frac{f'_c(1)}{2\pi} \mathcal{C}_\Sigma \right) \log \left(\frac{1}{\epsilon} \right) \quad (5.54)$$

$$= - \frac{c}{2\pi} \left(\mathcal{R}_\Sigma + \frac{11}{6} \mathcal{K}_\Sigma - \frac{17}{9} \mathcal{C}_\Sigma \right) \log \left(\frac{1}{\epsilon} \right). \quad (5.55)$$

The ratio $C_E^{\text{univ}}/S_{EE}^{\text{univ}}$ thus depends on a , c and the geometric quantities.

Similar expressions with integrals over local geometric quantities for the universal terms should exist for other even-dimensional CFTs [83]. For odd dimensions, this is not possible for the universal constant term, making the computation of the universal term more challenging.¹⁵

For spherical entangling surfaces of CFTs with gravity duals, we saw that there was a natural choice of UV regulation such that the entanglement entropy and the capacity of entanglement were proportional to each other. It would be interesting to see, whether deformed shapes would also have similarly natural choices for UV regulation.

VI. THE CAPACITY OF ENTANGLEMENT UNDER PERTURBATION WITH RELEVANT OPERATORS

In previous sections we studied conformal field theories, and in particular in $1+1$ dimensions we found that the leading terms of the first cumulants, the capacity of entanglement $\langle K^2 \rangle_c = C_E$ and the entanglement entropy $\langle K \rangle_c = S_{EE}$ were equal. In this section we will study what happens when the theory develops a mass gap or more generally is deformed by relevant operators to break the conformal invariance.

As a warm-up example, we study the anisotropic Heisenberg XY model. We find that the relationship $C_E = S_{EE}$ is broken, when the parameters are moved away from the critical domains, and the equivalent free fermion system develops a mass gap. We will also study how the divergence structure of the capacity of entanglement alters under perturbing away from criticality. The first order perturbation will contain a new universal \log^2 divergent term.

¹⁵For general dimensions, there has been interest in e.g., the universal term of entanglement entropy by corners and conical singularities [84–86] and small deformations [87]. Interestingly, with the inclusion of conical singularities in the entanglement surface, an additional log term emerges, modifying the form of universal terms for all Rényi entropies [86]

$$S_a^{\text{univ}} = \begin{cases} (-1)^{\frac{d-1}{2}} a_\alpha^{(d)}(\Omega) \log(R/\delta) & d \text{ odd} \\ (-1)^{\frac{d-2}{2}} a_\alpha^{(d)}(\Omega) \log^2(R/\delta) & d \text{ even} \end{cases} \quad (5.56)$$

where the positive $a_\alpha^{(d)}$ depend on the opening angle, $0 \leq \Omega \leq \pi$. These kinds of terms would then also appear in the capacity of entanglement.

After this concrete example, we will now perform a more general analysis and study CFTs perturbed with relevant operators. We will be considering two different entangling surfaces, planar and spherical. The strategy that we use follows the work [28,30].

A. The capacity of entanglement in the anisotropic Heisenberg XY spin chain

The anisotropic Heisenberg XY spin chain is a nice example of a system with a nontrivial phase structure. The system can be mapped to noninteracting fermions, with critical domains having a gapless spectrum, but elsewhere the spectrum is gapped. Using analytical results by Korepin *et al.* [33,47] on Rényi entropy in the anisotropic Heisenberg spin chain, we compare S_{EE} with the capacity of entanglement. We find that at criticality $C_E = S_{EE}$, but moving away from the critical domains causes deviations from this relation.

The Hamiltonian of the model is

$$H = - \sum_{j=-\infty}^{\infty} [(1+\gamma)\sigma_j^x\sigma_{j+1}^x + (1-\gamma)\sigma_j^y\sigma_{j+1}^y + h\sigma_j^z], \quad (6.1)$$

where γ parametrizes anisotropy, and h is the external magnetic field. The phase diagram of the model has three regions:

$$1a: 4(1-\gamma^2) < h^2 < 4$$

$$1b: h^2 < 4(1-\gamma^2)$$

$$2: h > 2$$

with the critical lines $\gamma = 0$, $h \leq 2$ where it becomes isotropic (the XX-model), and $h = h_c = 2$ corresponding to the critical value of the magnetic field. The line $h = h_f(\gamma) = 2\sqrt{1-\gamma^2}$ separating the regions 1a, b is not a phase transition, but the entanglement entropy has a weak singularity on it, and ground state is double degenerate with a basis given by two product states.

Korepin *et al.* computed the entanglement and Rényi entropies for a spin chain of N spins, considering the system in its ground state, separating a subsystem A of L spins, and considered the double scaling limit $N, L \rightarrow \infty$ with L/N fixed. For the Rényi entropy (with the reduced density matrix ρ_A and exponent α) they obtained

$$S_R(\rho_A, \alpha) = \begin{cases} \frac{1}{6} \frac{\alpha}{1-\alpha} \ln(kk') - \frac{1}{3} \frac{1}{1-\alpha} \ln\left(\frac{\theta_2(0, q^\alpha)\theta_4(0, q^\alpha)}{\theta_3^2(0, q^\alpha)}\right) + \frac{1}{3} \ln 2, & h > 2 \\ \frac{1}{6} \frac{\alpha}{1-\alpha} \ln\left(\frac{k'}{k^2}\right) - \frac{1}{3} \frac{1}{1-\alpha} \ln\left(\frac{\theta_3^2(0, q^\alpha)}{\theta_2(0, q^\alpha)\theta_4(0, q^\alpha)}\right) + \frac{1}{3} \ln 2, & h < 2 \end{cases} \quad (6.2)$$

with the elliptic modulus parameter

$$k = \begin{cases} \sqrt{(h/2)^2 + \gamma^2 - 1}/\gamma & : 1a \\ \sqrt{(1-h^2/4 - \gamma^2)/(1-h^2/4)} & : 1b \\ \gamma/\sqrt{(h/2)^2 + \gamma^2 - 1} & : 2 \end{cases} \quad (6.3)$$

and its complement $k' = \sqrt{1-k^2}$, and the nome

$$q = e^{-\pi I(k')/I(k)} \equiv e^{-\pi\tau_0} \quad (6.4)$$

where $I(k)$ is the complete elliptic integral of the first kind. Either by taking the limit $\alpha = 1$, or computing the first derivative of the generating function $\tilde{k}(\alpha) \equiv (1-\alpha)S_R(\rho_A, \alpha)$, the entanglement entropy becomes

$$S_{EE}(\rho_A) = \begin{cases} \frac{1}{6} \left[\ln \frac{4}{kk'} + (k^2 - k'^2) \frac{2I(k)I(k')}{\pi} \right], & h > 2 \\ \frac{1}{6} \left[\ln \left(\frac{4k^2}{k'} \right) + (2 - k^2) \frac{2I(k)I(k')}{\pi} \right], & h < 2. \end{cases} \quad (6.5)$$

The result for $h > 2$ comes from an expression

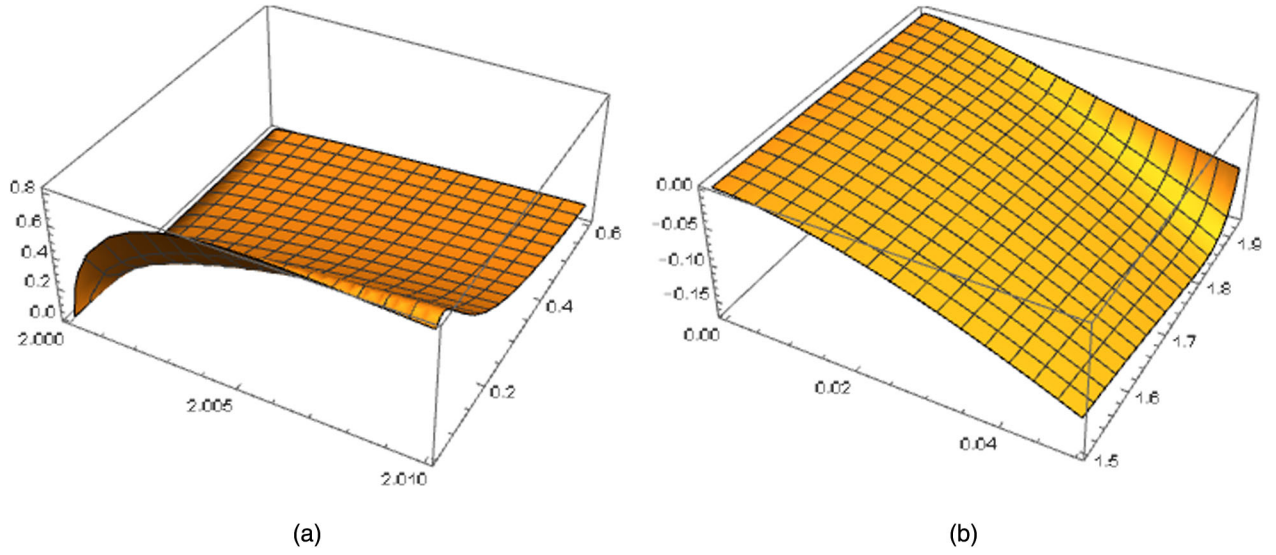
$$S_{EE}(\rho_A) = \frac{1}{6} \ln \frac{4}{kk'} + \frac{1}{12} \ln q + 2 \ln q \sum_{m=0}^{\infty} \frac{(2m+1)q^{2m+1}}{1+q^{2m+1}}, \quad (6.6)$$

where the latter series can be found from Abramowitz and Stegun in closed form with the elliptic integrals [88] to recover the result in the above. In region $h > 2$, the capacity of entanglement, from the double derivative of the generating function, becomes

$$C_E(\rho_A) = 2 \ln q \sum_{m=0}^{\infty} \left[\frac{(2m+1)q^{2m+1}}{1+q^{2m+1}} \right]^2, \quad (6.7)$$

We have not identified the closed form for the infinite series, but appears clear that in general the capacity of entanglement differs from S_{EE} . We can study this in more detail near the critical phases, to find that at criticality $C_E = S_{EE}$ while they begin to deviate moving away from the critical phase.

For $\gamma \neq 0$, near the phase boundary at $h_c = 2$, for the leading and next-to-leading order contributions we obtain

FIG. 5. Samples of the difference $S_{EE} - C_E$ near the critical line $h_c = 2$ (a) and near the critical line $\gamma = 0$ (b).

$$S_{EE}(\gamma, h) = \frac{1}{6} \log \left(\frac{16\gamma^2}{|h-2|} \right) + \left(\frac{5-\gamma^2-3\log(\frac{16\gamma^2}{|h-2|})}{24\gamma^2} \right) (h-2) + \mathcal{O}((h-2)^2) \quad (6.8)$$

$$C_E(\gamma, h) = \frac{1}{6} \log \left(\frac{16\gamma^2}{|h-2|} \right) + \left(\frac{2-\gamma^2-6\log(\frac{16\gamma^2}{|h-2|})+3\log^2(\frac{16\gamma^2}{|h-2|})}{24\gamma^2} \right) (h-2) + \mathcal{O}((h-2)^2), \quad (6.9)$$

where the expansion parameter is the inverse of the relevant length scale ξ ,

$$\xi^{-1} = |h-2|. \quad (6.10)$$

The leading term in the entanglement entropy matches with the result for a conformal field theory with $c = 1/2$,

$$S_{EE} = \frac{c}{3} \log \xi. \quad (6.11)$$

We can now investigate how the deviation from $C_E = S_{EE}$ happens in the vicinity of the critical line. The difference of the two, $S_{EE} - C_E$ is depicted in the Fig. 5(a). In this region S_{EE} grows faster than the capacity of entanglement C_E .

On the other hand, near the isotropic critical line $\gamma = 0$, for $h < 2$, we find

$$S_{EE}(\rho_A) = \frac{1}{6} \log \left(\frac{4(4-h^2)}{\gamma^2} \right) + \frac{(5+3\log(\frac{\gamma^2}{4(4-h^2)}))}{6(h^2-4)} \gamma^2 + \mathcal{O}(\gamma^4) \quad (6.12)$$

$$C_E(\rho_A) = \frac{1}{6} \log \left(\frac{4(4-h^2)}{\gamma^2} \right) - \frac{(2+3\log^2(\frac{\gamma^2}{16-4h^2})+6\log(\frac{\gamma^2}{16-4h^2}))}{6(4-h^2)} \gamma^2 + \mathcal{O}(\gamma^4), \quad (6.13)$$

with the relevant inverse length scale

$$\xi^{-1} = |\gamma|, \quad (6.14)$$

so the leading term in S_{EE} matches with the CFT result (6.11) with $c = 1$. Figure 5(b) depicts $S_{EE} - C_E$ in this case. In this region C_E grows faster than S_{EE} . The saddle seen in Fig. 5 likely reflects the boundary $h_f(\gamma) = 2\sqrt{1-\gamma^2}$ separating the phase regions 1a and 1b, but a full matching would require a higher order calculation than (6.8). From the approximate results it appears that S_{EE} is larger in the vicinity of the $h_c = 2$ critical line, while C_E is larger in the vicinity of the $\gamma = 0$ critical line.

B. Field theories perturbed with relevant operators

Much of the initial discussion here has already appeared in [29]. Let \mathcal{M} be a d -dimensional Euclidean manifold. Consider its ground state $|0\rangle$ and the corresponding density matrix $|0\rangle\langle 0|$. We partition the manifold to V and \bar{V} and take a trace of the density matrix over the d.o.f. in \bar{V} , obtaining the reduced density matrix in V , ρ_V . Inspired by the replica trick, we consider a cut \mathcal{C} and its both sides, \mathcal{C}_+ and \mathcal{C}_- . The density matrix ρ can be expressed as a path integral

$$[\rho_V]_{\phi_-\phi_+} = \frac{1}{\mathcal{N}} \int_{\substack{\phi(\mathcal{C}_-) = \phi_- \\ \phi(\mathcal{C}_+) = \phi_+}} e^{-I[\phi]}. \quad (6.15)$$

Now, we let I be an action with a small perturbation by operator O , i.e., $I[\phi] = I_0[\phi] - g \int_{\mathcal{M}} O$. Note that the action

governs the whole system, so the integral is over the whole manifold \mathcal{M} , not just V . In the limit of small g , we expand

$$[\rho_V]_{\phi_-\phi_+} = \frac{1}{\mathcal{N}} \int_{\substack{\phi(C_-)=\phi_- \\ \phi(C_+)=\phi_+}} e^{-I_0[\phi]} \left(1 + g \int_M O + \frac{g^2}{2} \int_M \int_M OO + \dots \right), \quad (6.16)$$

with

$$\mathcal{N} = \mathcal{N}_0 \left(1 + g \int_M \langle O \rangle_0 + \frac{g^2}{2} \int_M \int_M \langle OO \rangle_0 + \dots \right), \quad (6.17)$$

where $\langle \dots \rangle_0$ is the vacuum expectation value in the unperturbed theory. To first order,

$$[\rho_V]_{\phi_-\phi_+} \approx \frac{1}{\mathcal{N}_0} \int_{\substack{\phi(C_-)=\phi_- \\ \phi(C_+)=\phi_+}} e^{-I_0[\phi]} \left(1 + g \int_M (O - \langle O \rangle_0) \right) \quad (6.18)$$

$$= \left[\rho_{V,0} + g \int_M \text{Tr}_{\bar{V}} \rho_0 (O - \langle O \rangle_0) \right]_{\phi_-\phi_+}, \quad (6.19)$$

where ρ_0 and $\rho_{V,0}$ are the unperturbed vacuum density matrices of the whole system and the subsystem, respectively. The vacuum expectation value of O vanishes whenever O is a pseudoprimary operator. We assume this to be the case from now on. In addition, we focus on operators with scaling dimension $\Delta > \frac{d}{2}$ as the perturbative expansion of CFTs fail whenever $\Delta \leq \frac{d}{2}$ [89].

To compute the Rényi entropies, we need

$$\text{Tr} \rho_V^n = \text{Tr} \rho_{V,0}^n + ng \int_M \text{Tr} [\rho_{V,0}^{n-1} \text{Tr}_{\bar{V}} \rho_0 O] + \mathcal{O}(g^2). \quad (6.20)$$

The first order correction to the entanglement entropy is then

$$-g \partial_g \partial_n \text{Tr} \rho_V^n|_{n=1, g=0} = g \int_M \text{Tr} [K \text{Tr}_{\bar{V}} \rho_0 O], \quad (6.21)$$

and to the capacity of entanglement,

$$g \partial_g \partial_n^2 \log \text{Tr} \rho_V^n|_{n=1, g=0} = g \int_M \text{Tr} [(K^2 - 2K) \text{Tr}_{\bar{V}} \rho_0 O] - 2 \text{Tr} [K] \text{Tr} [K \text{Tr}_{\bar{V}} \rho_0 O]. \quad (6.22)$$

In the case of a QFT with a planar entangling surface or a CFT with a spherical entangling surface, K is known and expressible as an integral of the energy-momentum tensor.

As a final remark, the last term in the perturbation of variance can be ignored as its contributions will cancel with the normalization constants of the former terms.

C. Planar entangling surface

We consider first a planar entangling surface $\Sigma = \mathbb{R}^{d-2}$, at $x_1 = 0, x_0 = 0$. For any Cauchy surface A that ends at the entangling surface, the modular Hamiltonian is¹⁶

$$K = -2\pi \int_A d^{d-2}y dx_1 n^\mu T_{\mu\nu} \xi^\nu, \quad (6.23)$$

where x_0 and x_1 are the coordinates transverse to the entangling surface and y are the coordinates parallel to Σ . In addition, $\xi = x_1 \partial_{x_0} - x_0 \partial_{x_1}$ is a Killing vector that keeps the entangling surface invariant and n is the normal vector to the Cauchy surface.

The usual choice for the Cauchy surface is to set $x_0 = 0$ which yields

$$K = -2\pi \int_\Sigma d^{d-2}y \int_0^\infty dx_1 x_1 T_{00}, \quad (6.24)$$

Due to the freedom in choosing the Cauchy surface, we can also write K as

$$K = 2\pi \int_\Sigma d^{d-2}y \int_{-\infty}^0 dx_1 x_1 T_{00}. \quad (6.25)$$

In addition, we can also rewrite the integral of the O contribution in the $\langle KO \rangle$ term as [30]

$$\int_{\mathbb{R}^d} d^d x O \rightarrow 2\pi \int_\Sigma d^{d-2}y \int_0^\infty dx_1 x_1 O. \quad (6.26)$$

1. Massive free scalar field theory in four Euclidean dimensions

As in [28], we start with the free scalar field theory, with a mass term as a relevant deformation: $g = -\frac{m^2}{2}$ and $\mathcal{O}(x) = \phi^2(x)$. The stress tensor of the massless theory is

$$T_{\mu\nu}^0 = \partial_\mu \phi \partial_\nu \phi - \frac{1}{2} \delta_{\mu\nu} (\partial\phi)^2. \quad (6.27)$$

Using Wick contractions, we can compute the terms appearing in the variation of variance. The two-point functions of the scalar fields is

$$\langle \phi(x) \phi(y) \rangle = \frac{1}{(d-2)\Omega_{d-1}} \frac{1}{(x-y)^{d-2}}. \quad (6.28)$$

We point the reader to the Appendices for computational details.

¹⁶For the sign, we follow the conventions of [29].

When computing the integrals, we will run into UV or IR divergences. These emerge when integrating perpendicular to the entangling surface. We will regulate these with ε and $1/m$ respectively. Summing all the contributions, we get

$$\delta C_E = g\mathcal{A}_\Sigma \frac{(d-3)\pi^{\frac{3-d}{2}}\Gamma(\frac{d}{2}-1)^2}{2^{d-1}(d-2)(d-4)\Gamma(\frac{d+1}{2})}(\delta^{4-d} - m^{d-4}) \quad (6.29)$$

$$\stackrel{d=4}{=} -g\frac{\mathcal{A}_\Sigma}{12\pi}\log(\delta m) \quad (6.30)$$

so in $d = 4$ the first order correction to the capacity of entanglement ΔS_{EE}^2 is

$$\delta C_E \stackrel{d=4}{=} \frac{m^2}{24\pi}\mathcal{A}_\Sigma \log(m\delta). \quad (6.31)$$

In the above, \mathcal{A}_Σ is the area of the entangling surface.

Interestingly, the first order perturbation to entanglement entropy has the form

$$\delta S_{EE} = g\mathcal{A}_\Sigma \frac{\pi^{\frac{3-d}{2}}\Gamma(\frac{d}{2})^2}{2^{d-1}(d-2)(d-4)\Gamma(\frac{d+1}{2})}(\delta^{4-d} - m^{d-4}) \quad (6.32)$$

$$\stackrel{d=4}{=} -\frac{g\mathcal{A}_\Sigma}{12\pi}\log(m\delta). \quad (6.33)$$

Thus, when $d = 4$, the correction terms are equal for entanglement entropy and the capacity of entanglement.

2. A general CFT

The computation for a general CFT is more involved. We need the general forms of $\langle T_{\mu\nu}\mathcal{O} \rangle$ and $\langle T_{\mu\nu}T_{\alpha\beta}\mathcal{O} \rangle$ for any CFT. Due to the tracelessness and sourcelessness of the energy-momentum tensor, the two-point function automatically vanishes for any CFT. Thus, the entanglement entropy receives no correction at first order in the coupling g . This is not the case for the capacity of entanglement. The computation requires some effort but is, nevertheless, straightforward. We leave the computational details to Appendices and move on to the results.

We consider two special cases. In the following, a_3 is a multiplicative numerical factor appearing in the three point function. In most cases, we do not know any physical interpretation for it.

First, for $\Delta = 2$ and all d we get a logarithmic contribution. In this case, the general first order correction to the capacity of entanglement is

$$\delta C_E = ga_3 \frac{4(d-2)\pi^{d+1}\cos(\frac{\pi d}{2})\Gamma(3-d)\log(\delta\Lambda)}{d(d-2)(d-3)-4}\mathcal{A}_\Sigma, \quad (6.34)$$

where Λ is an IR regulator and a_3 is a numerical coefficient. Interestingly, the result has an IR divergence only for even $d \geq 4$, in agreement with the breakdown of CFT perturbations for $\Delta \leq \frac{d}{2}$. For $d = 2$, the correction vanishes. In fact, $\langle T_{\mu\nu}T_{\alpha\beta}\mathcal{O} \rangle = 0$ at $d = 2$, so just like S_{EE} , C_E receives no corrections at leading order, for all Δ for which the perturbative approach is valid.

Second, we consider the weight $\Delta = d - 2$ which corresponds e.g., to a mass term in a scalar field theory. The first order correction is

$$\delta C_E = ga_3 \frac{32(d-3)\pi^{d+1}}{d(d-2)(d-4)^2\Gamma(d-1)}(\Lambda^{d-4} - \delta^{4-d})\mathcal{A}_\Sigma. \quad (6.35)$$

Once again, we see the emergence of the divergences near $d = 4$. However, the divergence for $d = 2$ is not a real divergence as the three-point function is zero. The expression vanishes when $d = 3$.

The overall numerical factor a_3 , for a conformally coupled scalar field theory with a mass perturbation $g = -\frac{m^2}{2}$, has the value

$$a_3 = \frac{d}{2(d-1)^2\Omega_{d-1}^3} = \frac{d[\Gamma(\frac{d}{2})]^3}{16\pi^{\frac{3d}{2}}(d-1)^2}. \quad (6.36)$$

To summarize, in these special cases we find divergent corrections δC_E at leading order in g . The exceptions are $d = 3$, $\Delta = 1$ and $d = 2$, $\Delta > 1$, where $\delta C_E = 0$. In particular the equality $C_E = S_{EE}$ receives no corrections at leading order in g .

D. Spherical entangling surfaces of a CFT

There is a conformal mapping from the Rindler wedge to the causal diamond of a ball. Hence, for a spherical entangling surface of radius R in Euclidean spacetime, the modular Hamiltonian for a CFT is

$$K = -2\pi \int_{B'} n^\mu T_{\mu\nu} \xi^\nu, \quad (6.37)$$

where B' is any Cauchy surface that ends at the spherical entangling surface. Once again, n is the normal vector to the Cauchy surface and $\xi = \frac{1}{2R}(R^2 + x_0^2 - r^2)\partial_0 + \frac{x_0 r}{R}\partial_r$ is a conformal Killing vector that keeps the entangling surface at $x_0 = 0$, $r = R$ invariant. With the simple choice $B' = B$, the modular Hamiltonian is

$$K = -2\pi \int_{B(0,R)} d^{d-1}x \frac{R^2 - r^2}{2R} T_{00}(0, x). \quad (6.38)$$

When computing the effects of the perturbation, only the $\langle KKO \rangle$ term is found to be nonzero at first order. We can recycle some of the computations done for the planar entangling surface. We leave many of the details to the Appendices. The correction to the capacity of entanglement is

$$\begin{aligned} \delta C_E &= g \int_{\mathbb{R}^d} d^d x_3 \langle KKO(x_3) \rangle \\ &= g(2\pi)^2 \int_{|x_1| \leq R} d^{d-1}x_1 \int_{|x_2| \leq R} d^{d-1}x_2 \\ &\quad \times \int_{\mathbb{R}^d} d^d x_3 \frac{(R^2 - r_1^2)(R^2 - r_2^2)}{4R^2} \\ &\quad \times \langle T_{tt}(0, x_1) T_{tt}(0, x_2) O(x_3) \rangle. \end{aligned} \quad (6.39)$$

The integral over the x_3 coordinates is essentially the same as for the planar entangling surface done in (A 2).

After the integral over x_3 , we are left with

$$\delta C_E = ga_3(2\pi)^2 \pi^{d/2} f(d, \Delta) \frac{\Gamma(\frac{d-\Delta}{2})^2 \Gamma(\Delta - \frac{d}{2})}{\Gamma(\frac{\Delta}{2} + 2) \Gamma(d - \Delta)} I_d \quad (6.40)$$

where a_3 is the numerical factor that already appeared in the planar computations and f is a rational function and its specific value for arbitrary d and Δ is seen in (A44). The gamma functions capture the possible IR divergences for $\Delta \leq d/2$ (where the perturbative approach breaks down [89]). In addition, the relevant perturbations are for $\Delta \leq d$. What remains is the computation of the integral I_d ,

$$\begin{aligned} I_d &= \int_{|x_1| \leq R} d^{d-1}x_1 \int_{|x_2| \leq R} d^{d-1}x_2 \frac{(R^2 - r_1^2)(R^2 - r_2^2)}{4R^2} \\ &\quad \times \frac{1}{(x_1 - x_2)^{d+\Delta}}. \end{aligned} \quad (6.41)$$

Further details on evaluating this integral can be found in Appendix B.

The perturbation has a universal logarithmic term whenever Δ is an even number. It is

$$\delta C_E^{(\log)} = -ga_3 f(d, \Delta) \frac{\pi^{\frac{3d}{2}-1} 2^{2-\Delta} \Gamma(-\frac{\Delta-1}{2}) \Gamma(2\Delta - \frac{1}{2}) \Gamma(\frac{d-\Delta}{2})^2 \Gamma(\Delta - \frac{d}{2}) R^{d-\Delta}}{\Gamma(\frac{d-1}{2}) \Gamma(\frac{\Delta}{2} + 2)^2 \Gamma(2\Delta) \Gamma(d - \Delta) \Gamma(\frac{1}{2}(d - \Delta + 4))} \log(\epsilon). \quad (6.42)$$

As a specific example, consider $d = 4$, then

$$\delta C_{E,d=4}^{(\log)} = ga_3 \frac{2\pi^{11/2} \Delta(\Delta + 2) R^{4-\Delta} \Gamma(\Delta - 3) \Gamma(2\Delta - \frac{1}{2})}{(3(\frac{\Delta}{2} - 4) \frac{\Delta}{2} + 10) \Gamma(2\Delta) \Gamma(\frac{\Delta}{2} + 2)^2} \log(\epsilon). \quad (6.43)$$

For other values of Δ , there is a universal constant term, which is

$$\delta C_E^{(\text{const})} = ga_3 \frac{f(d, \Delta) 2^{8-d} \pi^{\frac{3d}{2}+1} \Gamma(1 - \frac{\Delta}{2}) \Gamma(\Delta - \frac{d}{2}) R^{d-\Delta}}{(\Delta^2 - 1)(d - \Delta)(d - \Delta + 2) \Gamma(\frac{d-1}{2}) \Gamma(\frac{\Delta}{2} + 2)^2 \Gamma(\frac{1}{2}(d - \Delta + 1))}. \quad (6.44)$$

Once again, we consider an example, $d = 4$, for which

$$\delta C_{E,d=4}^{(\text{const})} = ga_3 \frac{8\pi^{13/2} \Delta(\Delta + 2) \Gamma(1 - \frac{\Delta}{2}) \Gamma(\Delta - 2) R^{4-\Delta}}{(3(\Delta - 8) \Delta + 40) \Gamma(\frac{5}{2} - \frac{\Delta}{2}) \Gamma(\frac{\Delta}{2} + 2)^2}. \quad (6.45)$$

To summarize, we have isolated explicit expressions for universal perturbative corrections to C_E in $d = 4$. Similar results can be obtained in other dimensions from (6.40).

VII. DISCUSSION AND OUTLOOK

In this paper we discussed several aspects of capacity of entanglement C_E , which encodes a particular feature of the

set of eigenvalues of a reduced density matrix. It can roughly be thought of as the variance of the eigenvalue distribution, and can be obtained in a straightforward way from the (analytically continued) Rényi entropies. It may have interesting divergences across quantum phase transitions (in which case particular critical exponents may appear) but we have not studied this particular aspect in this paper.

If the reduced density matrix is viewed as a thermal density matrix, and we are allowed to change its “temperature” by considering a suitable one-parameter flow [the “modular flow” $\rho(\theta) = \rho^{1+\theta}/\text{Tr}(\rho^{1+\theta})$], then we can relate capacity of entanglement to the ordinary heat capacity defined with respect to this fiducial temperature. At the same time this allows us to connect to various information

theoretic quantities like fidelity susceptibility and Fisher information. Unfortunately, in an actual generic physical subsystem there is no natural operation corresponding to this change in “temperature” and therefore these relations remain somewhat academic.

While in general C_E can be much larger than S_{EE} , we found several situations where C_E is comparable to S_{EE} . First and foremost this happens in CFT’s with a holographic dual where both S_{EE} and C_E have an area law divergence. In principle the coefficients that appear in this divergence are scheme dependent in $d > 2$, but if the reduced density matrix is that associated to a ball in the ground state there is a fairly natural scheme in which S_{EE} is precisely equal to C_E .

We have studied various setups in order to gain some intuition for this approximate equality. First, following earlier work, we related the capacity of entanglement to fluctuations of $U(1)$ charges fluctuating in and out of the subregion, communicating entanglement. In this context, if the fluctuations are Gaussian distributed, or if the total particle number $N \rightarrow \infty$, we found that the capacity of entanglement becomes equal to the entanglement entropy, $C_E = S_{EE}$.

In a second setup, which we expect to be closely related to the previous one with fluctuation $U(1)$ charges, we considered random bipartite entanglement and observed that for randomly entangled pairs of qubits C_E and S_{EE} are also approximately equal.

All this then suggests that the approximate equality of C_E and S_{EE} implies that entanglement is effectively carried by randomly entangled UV pairs of qubits and not by large numbers of maximally entangled EPR pairs, as in the latter case C_E would be much smaller than S_{EE} . This in particular applies to the quasiparticle picture sometimes used to describe entanglement growth after quenches: if C_E is proportional to S_{EE} , the quasiparticles should be approximately randomly entangled and not in maximally entangled EPR pairs.

The area law for capacity of entanglement seems to extend quite generally to an area law for Rényi entropies. This is perhaps not that surprising, because a volume law would disagree with the basic observation that the Rényi entropy for a subregion equals that of the complement of the subregion. In addition, such an area law is also in agreement with the above picture of randomly entangled UV pairs of qubits, and as we explained in Sec. VC, in agreement with a local Rindler picture of the entangled d.o.f. This is all self-consistent, as in Rindler space the entanglement is thermally distributed and therefore also not mostly carried by maximally entangled EPR pairs.

In AdS/CFT, C_E has a relatively simply bulk interpretation, given by metric fluctuations integrated over the Ryu-Takayanagi surface. This relationship is strongly reminiscent of the relation between response functions and fluctuations in Landau-Ginzburg theories. It is not entirely

obvious where the equality $C_E \sim S_{EE}$ comes from in this computation. One would expect a divergence to arise if one or both points in the graviton propagator approach the boundary of the Ryu-Takayanagi surface. If one point approaches the boundary there is an area divergence from the integral over that boundary point, but that gets reduced by the graviton propagator which decays at long distances. Therefore the divergence must come from the region of integration where both points are close to the boundary.¹⁷ Some naive power counting suggests that this gives indeed the right behavior, but it would be nice to examine this in more detail and try to connect it to the previous discussion.

One of the motivations for this work was to study quantum fluctuations in the metric at the RT surface and to study the validity of the semiclassical approximation. Since most of the contributions seem to come from the region near the boundary, capacity as we have defined it is perhaps not a very good probe of the size of bulk metric fluctuations. A better probe would be to consider fluctuations in manifestly finite quantities such as mutual information and relative entropy. It is straightforward to find a generalization of C_E for the case of relative entropy.¹⁸ It is given by

$$C_E(\rho|\sigma) = \text{Tr}(\rho(\log \rho - \log \sigma)^2) - (\text{Tr}(\rho(\log \rho - \log \sigma)))^2 \quad (7.1)$$

and a candidate quantity that generalize C_E to mutual information is $C_E(\rho_{AB}|\rho_A \otimes \rho_B)$. It would be interesting to study these quantities in more detail but we leave that to future work.

The equality of C_E and S_{EE} for (suitably regularized) CFT’s with a holographic dual suggests that it may be possible to extract a necessary criterion for the existence of a holographic dual from consideration of capacity of entanglement. This led us to study what happens when one breaks conformal invariance and we found that the capacity of entanglement of the entanglement entropy starts to deviate from the entanglement entropy. As a demonstration, we considered the anisotropic Heisenberg XY spin chain, where the equality of C_E and S_{EE} holds at criticality but is broken as the parameters move from the critical lines. We also studied deformations of CFTs by relevant operators, and found similar results. In that case, the analysis is

¹⁷This argument concerns the leading divergence. The equality of C_E and S_{EE} in quenches (in 2d CFT) indicates that contributions from deeper in the bulk (where the RT surface crosses the collapsing shell) are also important.

¹⁸While this expression is fairly obvious as it stands, it can also be obtained from various one-parameter generalizations of the relative entropy called Rényi divergences [90–93] by differentiating and setting the parameter equal to 1, similar to what we did for capacity of entanglement. This connection may be helpful in order to e.g., establish monotonicity properties of $C_E(\rho|\sigma)$. In CFTs and their gravity duals, Rényi divergences were recently studied in [94].

complicated by various singularities whose precise understanding we leave to future work.

Finally, we notice that finding estimates of fluctuations in the reduced density matrix as we dial external parameters is also of great relevance in putting bounds on the accuracy of local measurement. We intend to explore this connection in more detail in the near future.

ACKNOWLEDGMENTS

J. J. and E. K. V. are in part supported by the Academy of Finland Grant No. 1297472. J. J. is also in part

supported by the University of Helsinki Graduate School PAPU, and E. K. V. is also in part supported by a grant from the Vilho, Yrjö and Kalle Väisälä Foundation. We thank U. Danielsson, N. Jokela, A. Kupiainen, A. Lawrence, D. Sarkar, W. Taylor, E. Tonni, T. Wiseman for useful discussions and comments while this work was in progress. J. d. B. and E. K. V. also thank the workshops “Quantum Gravity, String Theory, and Holography”, “Black Holes and Emergent Spacetime”, “Field theory anthropics and naturalness in cosmology,” and “Holography: What is going on these days?” for hospitality and partial support during this work.

APPENDIX A: COMPUTATIONAL DETAILS OF THE PERTURBATIVE COMPUTATIONS WITH PLANAR ENTANGLING SURFACE

1. Mass deformation of a massless free scalar field theory

First, we need

$$\langle \mathcal{O}(x) T_{\mu\nu}^0(z) \rangle = \frac{2(x-z)_\mu(x-z)_\nu - \delta_{\mu\nu}(x-z)^2}{\Omega_{d-1}^2(x-z)^{2d}}, \quad (\text{A1})$$

where $\Omega_{d-1} = \frac{2\pi^{\frac{d}{2}}}{\Gamma(\frac{d}{2})}$ is the surface area of the $d-1$ sphere and

$$\begin{aligned} \langle \mathcal{O}(x) T_{\mu\nu}^0(z) T_{\alpha\beta}^0(\bar{z}) \rangle &= \frac{2}{\Omega_{d-1}^3} \frac{1}{(x-z)^d(x-\bar{z})^d(z-\bar{z})^d} \\ &\times \left((x-z)_\mu(x-\bar{z})_\alpha \left[\delta_{\nu\beta} - d \frac{(z-\bar{z})_\beta(z-\bar{z})_\nu}{(z-\bar{z})^2} \right] + \text{permutations of } (\alpha, \beta, \mu, \nu) \right. \\ &+ \delta_{\alpha\beta} \delta_{\mu\nu} \left[(x-z) \cdot (x-\bar{z}) - d \frac{((x-\bar{z}) \cdot (z-\bar{z}))((x-z) \cdot (z-\bar{z}))}{(z-\bar{z})^2} \right] \\ &- \delta_{\mu\nu} \left[(x-\bar{z})_\alpha(x-z)_\beta - d \frac{(x-\bar{z})_\alpha(z-\bar{z})_\beta[(x-z) \cdot (z-\bar{z})]}{(z-\bar{z})^2} + (\alpha \leftrightarrow \beta) \right] \\ &\left. - \delta_{\alpha\beta} \left[(x-z)_\mu(x-\bar{z})_\nu - d \frac{(x-z)_\mu(z-\bar{z})_\nu[(x-z) \cdot (z-\bar{z})]}{(z-\bar{z})^2} + (\mu \leftrightarrow \nu) \right] \right). \quad (\text{A2}) \end{aligned}$$

When we put in $\alpha, \beta, \mu, \nu = 0$ and set the stress energy tensor at the same point in time, we get

$$\langle \mathcal{O}(x) T_{00}^0(z) T_{00}^0(\bar{z}) \rangle = \frac{2}{\Omega_{d-1}^3} \frac{1}{(x-z)^d(x-\bar{z})^d(z-\bar{z})^d} \left[(x-z) \cdot (x-\bar{z}) - d \frac{(x-\bar{z}) \cdot (z-\bar{z})((x-z) \cdot (z-\bar{z}))}{(z-\bar{z})^2} \right]. \quad (\text{A3})$$

We now compute the required integrals, following [28]. We begin with the two-point function as it is simpler. Using time translation symmetry, we set the integral to the form

$$\langle \mathcal{O}K \rangle = (2\pi)^2 \int_{\Sigma} d^{d-2}y \int_{\Sigma} d^{d-2}\bar{y} \int_0^\infty dx_1 \int_{-\infty}^0 d\bar{x}_1 \langle \mathcal{O}(x) T_{00}(\bar{x}) \rangle \quad (\text{A4})$$

$$= \int_{\Sigma} d^{d-2}y \int_{\Sigma} d^{d-2}\bar{y} \int_0^\infty dx_1 \int_{-\infty}^0 d\bar{x}_1 \frac{-x_1 \bar{x}_1 (2\pi)^2}{\Omega_{d-1}^2((x_1 - \bar{x}_1)^2 + (y - \bar{y})^2)^{d-1}}. \quad (\text{A5})$$

We first shift $y \rightarrow y + \bar{y}$ and compute the integral over y and \bar{y}

$$\langle OK \rangle = \frac{-(2\pi)^2 \Omega_{d-3} \sqrt{\pi} \Gamma[\frac{d}{2} - 1]}{2^{d-1} \Gamma[\frac{d}{2} - \frac{1}{2}] \Omega_{d-1}^2} \mathcal{A}_\Sigma \int_0^\infty dx_1 \times \int_0^\infty d\bar{x}_1 \frac{x_1 \bar{x}_1}{(x_1 - \bar{x}_1)^d} \quad (\text{A6})$$

$$= \frac{(2\pi)^2 \Omega_{d-3} \sqrt{\pi} \Gamma[\frac{d}{2} - 1]}{(d^2 - 3d + 2) 2^{d-1} \Gamma[\frac{d}{2} - \frac{1}{2}] \Omega_{d-1}^2} \times \mathcal{A}_\Sigma \int_0^\infty dx_1 \frac{1}{x_1^{d-3}}. \quad (\text{A7})$$

We compute the integral of x_1 only from δ to m^{-1} to regulate both the UV and IR divergences. We take the limit $d \rightarrow 4$, yielding

$$\mathcal{A}_\Sigma \frac{\pi^{\frac{3-d}{2}} \Gamma(\frac{d}{2})^2}{2^{d-1} (d-2)(d-4) \Gamma(\frac{d+1}{2})} (\delta^{4-d} - m^{d-4}) \stackrel{d \rightarrow 4}{=} \frac{-\mathcal{A}_\Sigma}{12\pi} \log(m\delta). \quad (\text{A8})$$

This computation also gives the first order correction to the entanglement entropy S_{EE} . In $d = 4$ the result is

$$\delta S_{EE} = \frac{m^2}{24\pi} \mathcal{A}_\Sigma \log(m\delta). \quad (\text{A9})$$

For the capacity of entanglement we also need the three-point function (A3),

$$\langle OKK \rangle = -(2\pi)^2 \int_\Sigma d^{d-2}y \int_\Sigma d^{d-2}\bar{y} \int_0^\infty dx_1 x_1 \times \int_{-\infty}^0 d\bar{x}_1 \bar{x}_1 \int d^d z \langle O(z) T_{00}(x) T_{00}(\bar{x}) \rangle \quad (\text{A10})$$

The Appendix in [28] provides us with a good guide through the calculation. One identity that will prove useful is

$$\int d^d x_3 \frac{1}{x_{23}^\gamma x_{13}^\beta} = \pi^{d/2} \frac{\Gamma(\frac{d-\gamma}{2}) \Gamma(\frac{d-\beta}{2}) \Gamma(\frac{\beta+\gamma-d}{2})}{\Gamma(\gamma/2) \Gamma(\beta/2) \Gamma(d - \frac{\beta+\gamma}{2})} x_{12}^{d-\beta-\gamma}, \quad (\text{A11})$$

where $x_{ij} = x_i - x_j$.

We first compute

$$\int_\Sigma d^{d-2}y \int_\Sigma d^{d-2}\bar{y} \int_0^\infty dx_1 x_1 \int_{-\infty}^0 d\bar{x}_1 \bar{x}_1 \times \int d^d z \frac{(x-z)_\rho (\bar{x}-z)^\rho}{(x-z)^d (\bar{x}-z)^d (x-\bar{x})^d}. \quad (\text{A12})$$

We can write the integrand in the form

$$\frac{(x-z)_\rho (\bar{x}-z)^\rho}{(x-z)^d (\bar{x}-z)^d (x-\bar{x})^d} = \frac{1}{(x-\bar{x})^d} \frac{\partial_{\rho,x} \partial_{\bar{x}}^\rho}{(d-2)^2} \left(\frac{1}{(z-x)^{d-2} (z-\bar{x})^{d-2}} \right), \quad (\text{A13})$$

which we can easily integrate over z to get

$$\frac{4\pi^{d/2}}{\Gamma(\frac{d-2}{2})(d-2)^2} \frac{1}{(x-\bar{x})^{2d-2}}. \quad (\text{A14})$$

The remaining integrals can be computed as before. Thus, the contribution from the first term in (A3) sums up to

$$\frac{2\pi^{1-\frac{d}{2}} \Gamma(\frac{d}{2})^3}{(d-4)(d-2)^2 \Gamma(d)} \mathcal{A}_\Sigma (\delta^{4-d} - m^{d-4}) \quad (\text{A15})$$

$$\stackrel{d \rightarrow 4}{=} -\frac{\mathcal{A}_\Sigma}{12\pi} \log(m\delta). \quad (\text{A16})$$

The contribution from the second term in (A3) is more involved. We rewrite the integrand

$$\frac{(z-\bar{x})_\rho (x-\bar{x})^\rho (z-x)_\lambda (x-\bar{x})^\lambda}{(z-x)^d (z-\bar{x})^d (x-\bar{x})^{d+2}} \quad (\text{A17})$$

$$= \frac{(x-\bar{x})^\rho (x-\bar{x})^\lambda}{(x-\bar{x})^{d+2}} \frac{\partial_{\rho,x} \partial_{\lambda,\bar{x}}}{(d-2)^2} \times \left(\frac{1}{(z-\bar{x})^{d-2} (z-x)^{d-2}} \right), \quad (\text{A18})$$

which we integrate over z to get

$$-\frac{2\pi^{d/2} (d-3)}{\Gamma(\frac{d-2}{2})(d-2)^2} \frac{1}{(x-\bar{x})^{2d-2}}, \quad (\text{A19})$$

which we know how to integrate with respect to the two other coordinates. The contribution from the second term sums up to

$$\frac{(d-3) d \pi^{1-\frac{d}{2}} \Gamma(\frac{d}{2})^3}{(d-4)(d-2)^2 \Gamma(d)} \mathcal{A}_\Sigma (\delta^{4-d} - m^{d-4}) \quad (\text{A20})$$

$$\stackrel{d \rightarrow 4}{=} -\frac{\mathcal{A}_\Sigma}{6\pi} \log(m\delta). \quad (\text{A21})$$

This gives the contribution of the three point function as

$$\langle KKO \rangle = \frac{(d-1) \pi^{1-\frac{d}{2}} \Gamma(\frac{d}{2})^3}{(d^2 - 6d + 8) \Gamma(d)} \mathcal{A}_\Sigma (\delta^{4-d} - m^{d-4}) \quad (\text{A22})$$

$$= -\frac{\mathcal{A}_\Sigma}{4} \log(m\delta). \quad (\text{A23})$$

Summing all the contributions we get

$$\langle KK\mathcal{O} \rangle - 2\langle K\mathcal{O} \rangle = \mathcal{A}_\Sigma \frac{(d-3)(d-2)\pi^{\frac{3-d}{2}}\Gamma(\frac{d}{2}-1)^2}{2^{d+1}(d-4)\Gamma(\frac{d+1}{2})} \times (\delta^{4-d} - m^{d-4}) \quad (\text{A24})$$

$$\stackrel{d=4}{=} -\frac{\mathcal{A}_\Sigma}{12\pi} \log(\delta m) \quad (\text{A25})$$

so in $d=4$ the first order correction to the capacity of entanglement ΔS_{EE}^2 is

$$\delta \Delta S_{EE}^2 \stackrel{d=4}{=} \frac{m^2}{24\pi} \mathcal{A}_\Sigma \log(m\delta). \quad (\text{A26})$$

Thus, under the mass deformation the correction to the capacity of entanglement grows as fast as the correction to the entanglement entropy.

2. A general CFT

The three-point function does not vanish in general and is quite complex [95]:

$$\langle T_{\mu\nu}(x_1) T_{\alpha\beta}(x_2) \mathcal{O}(x_3) \rangle = \frac{I_{\mu\nu,\lambda\kappa}(x_{13}) I_{\alpha\beta,\sigma\rho}(x_{23}) t_{\lambda\kappa\sigma\rho}(X_{12})}{(x_{12})^{2d-\Delta} (x_{13})^\Delta (x_{23})^\Delta}, \quad (\text{A27})$$

where

$$t_{\lambda\kappa\sigma\rho} = a_1 h_{\lambda\kappa}^1(\hat{X}) h_{\sigma\rho}^1(\hat{X}) + a_2 h_{\lambda\kappa\sigma\rho}^2(\hat{X}) + a_3 h_{\lambda\kappa\sigma\rho}^3, \quad (\text{A28})$$

$$h_{\alpha\beta}^1(\hat{X}) = \hat{X}_\alpha \hat{X}_\beta - \frac{1}{d} \delta_{\alpha\beta} \quad (\text{A29})$$

$$h_{\lambda\kappa\sigma\rho}^2(\hat{X}) = \hat{X}_\lambda \hat{X}_\sigma \delta_{\kappa\rho} + (\lambda \leftrightarrow \kappa, \sigma \leftrightarrow \rho) \quad (\text{A30})$$

$$- \frac{4}{d} \hat{X}_\lambda \hat{X}_\kappa \delta_{\sigma\rho} - \frac{4}{d} \hat{X}_\sigma \hat{X}_\rho \delta_{\lambda\kappa} + \frac{4}{d^2} \delta_{\alpha\beta} \delta_{\sigma\rho}, \quad (\text{A31})$$

$$h_{\lambda\kappa\sigma\rho}^3 = \delta_{\lambda\sigma} \delta_{\kappa\rho} + \delta_{\lambda\rho} \delta_{\kappa\sigma} - \frac{2}{d} \delta_{\lambda\kappa} \delta_{\sigma\rho}, \quad (\text{A32})$$

$$I_{\mu\nu,\lambda\kappa}(x) = \frac{1}{2} (I_{\mu\lambda}(x) I_{\nu\kappa}(x) + I_{\mu\kappa}(x) I_{\nu\lambda}(x)) - \frac{1}{d} \delta_{\mu\nu} \delta_{\lambda\kappa}, \quad (\text{A33})$$

$$I_{\alpha\beta}(x) = \delta_{\alpha\beta} - 2 \frac{x_\alpha x_\beta}{x^2}, \quad (\text{A34})$$

$$\hat{X}_\mu = \frac{X_\mu}{\sqrt{X^2}}, \quad (\text{A35})$$

$$X_{12} = -X_{21} = \frac{x_{13}}{x_{13}^2} - \frac{x_{23}}{x_{23}^2}. \quad (\text{A36})$$

The condition of conservation of momentum and energy lead to two conditions for the coefficients a_1 , a_2 and a_3 ,

$$a_1 + 4a_2 - \frac{1}{2} (d-\Delta)(d-1)(a_1 + 4a_2) - d\Delta a_2 = 0 \quad (\text{A37})$$

$$a_1 + 4a_2 + d(d-\Delta)a_2 + d(2d-\Delta)a_3 = 0. \quad (\text{A38})$$

Thus, the three-point function is defined up to a multiplicative constant. Next, we use a few helpful relations [95]:

$$\begin{aligned} I_{\mu\alpha}(x_1 - x_3) X_{12,\alpha} &= -\frac{(x_1 - x_2)^2}{(x_3 - x_2)^2} X_{23,\mu}, \\ I_{\mu\alpha}(x_2 - x_3) X_{12,\alpha} &= -\frac{(x_1 - x_2)^2}{(x_1 - x_3)^2} X_{31,\mu}, \\ I_{\mu\alpha}(x_1 - x_3) I_{\alpha\nu}(x_3 - x_2) &= I_{\mu\nu}(x_1 - x_2) + 2(x_1 - x_2)^2 \\ &\quad \times X_{23,\mu} X_{31,\nu}. \end{aligned} \quad (\text{A39})$$

The a_1 dependent term becomes (sans the factors of $|x_i - x_j|$ in the denominator)

$$h_{\mu\nu}^1(\hat{X}_{23}) h_{\alpha\beta}^1(\hat{X}_{31}). \quad (\text{A40})$$

The a_2 dependent term becomes (the permutations produce 3 additional terms)

$$\begin{aligned} &(\hat{X}_{23,\mu} \hat{X}_{31,\alpha} (I_{\nu\beta}(x_{12}) + 2\hat{X}_{23,\beta} \hat{X}_{31,\nu}) + (\mu \leftrightarrow \nu, \alpha \leftrightarrow \beta)) \\ &+ \frac{4}{d} \left(\frac{\delta_{\mu\nu} \delta_{\alpha\beta}}{d} - \hat{X}_{23,\mu} \hat{X}_{23,\nu} \delta_{\alpha\beta} - \hat{X}_{31,\alpha} \hat{X}_{31,\beta} \delta_{\mu\nu} \right). \end{aligned} \quad (\text{A41})$$

The a_3 dependent term becomes

$$\begin{aligned} &I_{\mu\alpha}(x_{12}) I_{\nu\beta}(x_{12}) + I_{\mu\beta}(x_{12}) I_{\nu\alpha}(x_{12}) + 8\hat{X}_{13,\alpha} \hat{X}_{13,\beta} \hat{X}_{23,\alpha} \hat{X}_{23,\beta} \\ &+ 2(I_{\mu\alpha}(x_{12}) \hat{X}_{31,\beta} \hat{X}_{23,\nu} + (\mu \leftrightarrow \nu, \alpha \leftrightarrow \beta)) - \frac{2}{d} \delta_{\mu\nu} \delta_{\alpha\beta}. \end{aligned} \quad (\text{A42})$$

The integrals can be computed in a similar manner as before and are straightforward. In this paper, we are only interested in the case where all spacetime indices are set to 0 and $x_{1,0} = x_{2,0}$.

After integrating over x_3 , we have the intermediate result

$$\begin{aligned} &g \int_{\mathbb{R}^d} d^d x_3 \langle KKO(x_3) \rangle \\ &= -ga_3 (2\pi)^2 \frac{\pi^{d/2} f(d, \Delta) \Gamma(\frac{d-\eta}{2})^2 \Gamma(\eta - \frac{d}{2})}{\Gamma(\frac{\eta}{2} + 2)^2 \Gamma(d - \eta)} \\ &\quad \times \int_0^\infty dx_1 \int_\Sigma d^{d-2} y \int_{-\infty}^0 d\bar{x}_1 \int_\Sigma d^{d-2} \bar{y} \frac{x_1 \bar{x}_1}{(x - \bar{x})^{d+\Delta}} \end{aligned} \quad (\text{A43})$$

where

$$f(d, \Delta) = \frac{(d-2)(\Delta-1)\Delta(\Delta+1)(\Delta+2)(d-\Delta)(d-\Delta+2)}{8((d-1)\Delta^2 - 2(d-1)d\Delta + (d-2)d(d+1))}. \quad (\text{A44})$$

APPENDIX B: MORE DETAILED COMPUTATIONS OF THE THREE-POINT FUNCTION INTEGRAL WITH SPHERICAL ENTANGLING SURFACE

In this Appendix, we try to give more details on the computation of (6.41) for the values appearing in the text.

We make use of the Fourier transformation identity

$$\frac{1}{x^{d+\Delta}} = \frac{\Gamma(\frac{-\Delta-1}{2})}{2^{d+\Delta}\sqrt{\pi}^{d-1}\Gamma(\frac{d+\Delta}{2})} \int d^{d-1}k \frac{e^{-i\vec{k}\cdot\vec{x}}}{k^{-\Delta-1}} \quad (\text{B1})$$

to separate the two integrals over the balls. The integrals over the angular coordinates of x_1 and x_2 contribute

$$\begin{aligned} \Omega_{d-3} \int_0^\pi d\theta \sin^{d-3} \theta e^{ikr \cos \theta} \\ = \Omega_{d-3} \sqrt{\pi} \left(\frac{2}{kr}\right)^{\frac{d-3}{2}} \Gamma\left(\frac{d}{2} - 1\right) J_{\frac{d-3}{2}}(kr). \end{aligned} \quad (\text{B2})$$

After also integrating over the angular coordinates of k , we have

$$\begin{aligned} \Omega_{d-3}^2 \Omega_{d-2} \frac{\pi^{\frac{3-d}{2}} \Gamma(\frac{d}{2} - 1)^2 \Gamma(\frac{-\Delta-1}{2})}{2^{\Delta+5} R^2 \Gamma(\frac{d+\Delta}{2})} \int_0^\infty dk \int_0^R dr_1 \int_0^R dr_2 \\ \times \frac{(R^2 - r_1^2)(R^2 - r_2^2) J_{\frac{d-3}{2}}(kr_1) J_{\frac{d-3}{2}}(kr_2)}{k^{-\Delta-2} (r_1 r_2)^{\frac{1-d}{2}}} \\ = \Omega_{d-3}^2 \Omega_{d-2} \frac{\pi^{\frac{3-d}{2}} \Gamma(\frac{d}{2} - 1)^2 \Gamma(\frac{-\Delta-1}{2})}{2^{\Delta+3} R^{\Delta-d} \Gamma(\frac{d+\Delta}{2})} \int_0^\infty dk \frac{(J_{\frac{d+1}{2}}(k))^2}{k^{\Delta+2}}, \end{aligned}$$

where we have removed the R dependence from the integral. We can compute the remaining integral explicitly,

$$I_{d,\Delta} = \int_0^\infty dk \frac{(J_{\frac{d+1}{2}}(k))^2}{k^{\Delta+2}} = \frac{\pi^{d-\frac{3}{2}} 2^{3-\Delta} \Gamma(1 - \frac{\Delta}{2}) R^{d-\Delta}}{(\Delta^2 - 1) \Gamma(\frac{d-1}{2}) \Gamma(\frac{1}{2}(d - \Delta + 4))} \quad (\text{B3})$$

It is still divergent for even values of Δ but is now finite for odd values of Δ . The divergences with even Δ emerge from the UV limit $k \rightarrow \infty$, which we will regulate with $k \rightarrow \frac{1}{\epsilon}$. The logarithmically divergent term is

$$I_{d,\text{even}\Delta}^{(\log)} = - \frac{\pi^{d-\frac{5}{2}} 2^{-\Delta} \Gamma(-\frac{\Delta}{2} - \frac{1}{2}) \Gamma(2\Delta - \frac{1}{2}) R^{d-\Delta}}{\Gamma(\frac{d-1}{2}) \Gamma(2\Delta) \Gamma(\frac{1}{2}(d - \Delta + 4))} \log(\epsilon). \quad (\text{B4})$$

-
- [1] N. Laflorencie, *Phys. Rep.* **646**, 1 (2016).
 - [2] S. Ryu and T. Takayanagi, *Phys. Rev. Lett.* **96**, 181602 (2006).
 - [3] M. Rangamani and T. Takayanagi, *Lect. Notes Phys.* **931**, 1 (2017).
 - [4] H. Yao and X.-L. Qi, *Phys. Rev. Lett.* **105**, 080501 (2010).
 - [5] J. Schliemann, *Phys. Rev. B* **83**, 115322 (2011).
 - [6] Y. Nakaguchi and T. Nishioka, *J. High Energy Phys.* **12** (2016) 129.
 - [7] X. Dong, *Nat. Commun.* **7**, 12472 (2016).
 - [8] X. Dong, A. Lewkowycz, and M. Rangamani, *J. High Energy Phys.* **11** (2016) 028.
 - [9] H. Li and F. D. M. Haldane, *Phys. Rev. Lett.* **101**, 010504 (2008).
 - [10] P. Calabrese and A. Lefevre, *Phys. Rev. A* **78**, 032329 (2008).
 - [11] A. M. Lauchli, [arXiv:1303.0741](https://arxiv.org/abs/1303.0741).
 - [12] R. Lundgren, J. Blair, P. Laurell, N. Regnault, G. A. Fiete, M. Greiter, and R. Thomale, *Phys. Rev. B* **94**, 081112 (2016).
 - [13] G. Young Cho, A. W. W. Ludwig, and S. Ryu, *Phys. Rev. B* **95**, 115122 (2017).
 - [14] A. Chandran, V. Khemani, and S. L. Sondhi, *Phys. Rev. Lett.* **113**, 060501 (2014).
 - [15] Y. O. Nakagawa and S. Furukawa, *Phys. Rev. B* **96**, 205108 (2017).
 - [16] V. Alba, P. Calabrese, and E. Tonni, *J. Phys. A* **51**, 024001 (2018).
 - [17] M. Alishahiha and A. Faraji Astaneh, *Phys. Rev. D* **96**, 086004 (2017).
 - [18] S. Banerjee, J. Erdmenger, and D. Sarkar, *J. High Energy Phys.* **08** (2018) 001.
 - [19] I. Klich, G. Refael, and A. Silva, *Phys. Rev. A* **74**, 032306 (2006).
 - [20] I. Klich and L. Levitov, *Phys. Rev. Lett.* **102**, 100502 (2009).
 - [21] H. F. Song, S. Rachel, and K. Le Hur, *Phys. Rev. B* **82**, 012405 (2010).
 - [22] H. F. Song, C. Flindt, S. Rachel, I. Klich, and K. Le Hur, *Phys. Rev. B* **83**, 161408R (2011).

- [23] H. F. Song, S. Rachel, C. Flindt, I. Klich, N. Laflorencie, and K. Le Hur, *Phys. Rev. B* **85**, 035409 (2012).
- [24] E. Perlmutter, *J. High Energy Phys.* **03** (2014) 117.
- [25] L. Y. Hung, R. C. Myers, and M. Smolkin, *J. High Energy Phys.* **10** (2014) 178.
- [26] D. M. Hofman and J. Maldacena, *J. High Energy Phys.* **05** (2008) 012.
- [27] D. M. Hofman, D. Li, D. Meltzer, D. Poland, and F. Rejon-Barrera, *J. High Energy Phys.* **06** (2016) 111.
- [28] V. Rosenhaus and M. Smolkin, *J. High Energy Phys.* **02** (2015) 015.
- [29] V. Rosenhaus and M. Smolkin, *J. High Energy Phys.* **12** (2014) 179.
- [30] V. Rosenhaus and M. Smolkin, *J. High Energy Phys.* **09** (2014) 119.
- [31] F. Schlögl, *J. Stat. Phys.* **46**, 135 (1987).
- [32] P. Grassberger and I. Procaccia, *Phys. Rev. Lett.* **50**, 346 (1983).
- [33] F. Franchini, A. R. Its, and V. E. Korepin, *J. Phys. A* **41**, 025302 (2008).
- [34] S.-J. Gu, *Int. J. Mod. Phys. B* **24**, 4371 (2010).
- [35] A. De Pasquale, D. Rossini, R. Fazio, and V. Giovannetti, *Nat. Commun.* **7**, 12782 (2016).
- [36] D. Braun, G. Adesso, F. Benatti, R. Floreanini, U. Marzolino, M. W. Mitchell, and S. Pirandola, *Rev. Mod. Phys.* **90**, 035006 (2018).
- [37] A. Uhlmann, *Rep. Math. Phys.* **9**, 273 (1976); R. Josza, *J. Mod. Opt.* **41**, 2315 (1994).
- [38] S. L. Braunstein and C. M. Caves, *Phys. Rev. Lett.* **72**, 3439 (1994); S. L. Braunstein, C. M. Caves, and G. J. Milburn, *Ann. Phys. (N.Y.)* **247**, 135 (1996); H. Cramer, *Mathematical Methods of Statistics* (Princeton University Press, Princeton, NJ, 1946).
- [39] G. Ruppeiner, *Phys. Rev. A* **20**, 1608 (1979); *Rev. Mod. Phys.* **67**, 605 (1995); D. Brody and N. Rivier, *Phys. Rev. E* **51**, 1006 (1995); P. Zanardi, P. Giorda, and M. Cozzini, *Phys. Rev. Lett.* **99**, 100603 (2007).
- [40] H. Casini, M. Huerta, and R. C. Myers, *J. High Energy Phys.* **05** (2011) 036.
- [41] L. Y. Hung, R. C. Myers, M. Smolkin, and A. Yale, *J. High Energy Phys.* **12** (2011) 047.
- [42] R. Brustein and A. Yarom, *Phys. Rev. D* **69**, 064013 (2004).
- [43] T. Popoviciu, *Mathematica* **9**, 129 (1935).
- [44] P. Calabrese, M. Mintchev, and E. Vicari, *Europhys. Lett.* **98**, 20003 (2012).
- [45] S. Ghosh, R. M. Soni, and S. P. Trivedi, *J. High Energy Phys.* **09** (2015) 069.
- [46] R. M. Soni and S. P. Trivedi, *J. High Energy Phys.* **01** (2016) 136.
- [47] F. Franchini, A. R. Its, V. E. Korepin, and L. A. Takhtajan, *Quantum Inf. Process.* **10**, 325 (2011).
- [48] J. de Boer, F. M. Haehl, M. P. Heller, and R. C. Myers, *J. High Energy Phys.* **08** (2016) 162.
- [49] B. Czech, L. Lamprou, S. McCandlish, B. Mosk, and J. Sully, *J. High Energy Phys.* **07** (2016) 129.
- [50] M. Miyaji, T. Numasawa, N. Shiba, T. Takayanagi, and K. Watanabe, *Phys. Rev. Lett.* **115**, 261602 (2015).
- [51] N. Lashkari and M. Van Raamsdonk, *J. High Energy Phys.* **04** (2016) 153.
- [52] E. B. Fel'dman and M. A. Yurishchev, *JETP Lett.* **90**, 70 (2009).
- [53] S. Majumdar, *arXiv:1005.4515*.
- [54] S. Lloyd and H. Pagels, *Ann. Phys. (N.Y.)* **188**, 186 (1988).
- [55] K. Zyczkowski and H.-J. Sommers, *J. Phys. A* **34**, 7111 (2001).
- [56] O. Giraud, *J. Phys. A* **40**, 2793 (2007).
- [57] D. N. Page, *Phys. Rev. Lett.* **71**, 1291 (1993).
- [58] S. K. Foong and S. Kanno, *Phys. Rev. Lett.* **72**, 1148 (1994); J. Sánchez-Ruiz, *Phys. Rev. E* **52**, 5653 (1995); S. Sen, *Phys. Rev. Lett.* **77**, 1 (1996).
- [59] C. Nadal, S. N. Majumdar, and M. Vergassola, *J. Stat. Phys.* **142**, 403 (2011).
- [60] X. Chen and A. W. W. Ludwig, *Phys. Rev. B* **98**, 064309 (2018).
- [61] V. A. Marchenko and L. A. Pastur, *Math. USSR Sb. N.S.* **1**, 457 (1967).
- [62] L. U. Ancarani and G. Gasaneo, *J. Phys. A* **42**, 395208 (2009).
- [63] X. Wen, S. Ryu, and A. W. W. Ludwig, *J. Stat. Mech.* (2018) 113103.
- [64] P. Calabrese and J. Cardy, *J. Stat. Mech.* (2016) 064003.
- [65] J. Cardy and E. Tonni, *J. Stat. Mech.* (2016) 123103.
- [66] P. Calabrese and J. Cardy, *J. Phys. A* **42**, 504005 (2009).
- [67] H. Casini, C. D. Fosco, and M. Huerta, *J. Stat. Mech.* (2005) P07007.
- [68] H. Casini and M. Huerta, *J. High Energy Phys.* **03** (2009) 048.
- [69] H. Casini and M. Huerta, *Classical Quantum Gravity* **26**, 185005 (2009).
- [70] N. D. Birrell and P. C. W. Davies, *Quantum Fields in Curved Space* (Cambridge University Press, Cambridge, England, 1982).
- [71] M. M. Wolf, *Phys. Rev. Lett.* **96**, 010404 (2006).
- [72] D. Gioev and I. Klich, *Phys. Rev. Lett.* **96**, 100503 (2006).
- [73] B. Swingle, *arXiv:1010.4038*.
- [74] W. Ding, A. Seidel, and K. Yang, *Phys. Rev. X* **2**, 011012 (2012).
- [75] B. Swingle, *Phys. Rev. Lett.* **105**, 050502 (2010).
- [76] M. Srednicki, *Phys. Rev. Lett.* **71**, 666 (1993).
- [77] H. Casini and M. Huerta, *J. Phys. A* **42**, 504007 (2009).
- [78] N. Kim, *Phys. Lett. B* **733**, 233 (2014).
- [79] S. N. Solodukhin, *Living Rev. Relativity* **14**, 8 (2011).
- [80] I. R. Klebanov, S. S. Pufu, S. Sachdev, and B. R. Safdi, *J. High Energy Phys.* **04** (2012) 074.
- [81] D. V. Fursaev, *J. High Energy Phys.* **05** (2012) 080.
- [82] X. Dong, *Phys. Rev. Lett.* **116**, 251602 (2016).
- [83] A. Lewkowycz and E. Perlmutter, *J. High Energy Phys.* **01** (2015) 080.
- [84] P. Bueno, R. C. Myers, and W. Witczak-Krempa, *Phys. Rev. Lett.* **115**, 021602 (2015).
- [85] P. Bueno, R. C. Myers, and W. Witczak-Krempa, *J. High Energy Phys.* **09** (2015) 091.
- [86] P. Bueno and R. C. Myers, *J. High Energy Phys.* **12** (2015) 168.

- [87] L. Bianchi, S. Chapman, X. Dong, D. A. Galante, M. Meineri, and R. C. Myers, *J. High Energy Phys.* **11** (2016) 180.
- [88] I. Peschel, *J. Stat. Mech.* (2004) P12005.
- [89] T. Nishioka, *Phys. Rev. D* **90**, 045006 (2014).
- [90] D. Petz, *Rep. Math. Phys.* **23**, 57 (1986).
- [91] F. Hiai, M. Mosonyi, D. Petz, and C. Beny, *Rev. Math. Phys.* **23**, 691 (2011).
- [92] M. Müller-Lennert, F. Dupuis, O. Szehr, S. Fehr, and M. Tomamichel, *J. Math. Phys. (N.Y.)* **54**, 122203 (2013).
- [93] M. M. Wilde, A. Winter, and D. Yang, *Commun. Math. Phys.* **331**, 593 (2014).
- [94] A. Bernamonti, F. Galli, R. C. Myers, and J. Oppenheim, *J. High Energy Phys.* **07** (2018) 111.
- [95] H. Osborn and A. C. Petkou, *Ann. Phys. (N.Y.)* **231**, 311 (1994).